STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 142420

To: Janet Epps-Ford **Location:** REM/2C05/2C18 **Art Unit:** 1635 Tuesday, January 18, 2005 **Case Serial Number:** 09/733306	**From:** Beverly Shears **Location:** Remsen Bldg. RM 1A54 **Phone:** 571-272-2528 beverly.shears@uspto.gov

Search Notes

STIC-Biotech/ChemLib

From: Epps-Ford, Janet
Sent: Wednesday, January 12, 2005 12:45 PM
To: STIC-Biotech/ChemLib
Subject: Protein Sequence Search

Please search SEQ ID NO: 4 of application 09/733,306 in all commercial amino acid databases and all published us patent amino acid databases.

4-312 am

****
STAFF USE ONLY

Searcher:_____
Searcher Phone: 2-
Date Searcher Picked up:____
Date Completed:_____
Searcher Prep/Rev. Time:___
Online Time:_____

****
Type of Search
NA Sequence: #_____
AA Sequence :#_____
Structure: #_____
Bibliographic:_____
Litigation:_____
Patent Family:_____
Other:_____

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other(Specify):_____

Date completed: _____
Searcher: Beverly e2528
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site
_____ STIC
_____ CM-1
_____ Pre-S

Type of Search
_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors
_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ Other CGN

PTO-1590 (9-90)

Epps-Ford
09/733306
seq. ID 4 w/interf

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2005, 15:40:37 ; Search time 156 Seconds
(without alignments)
717.459 Million cell updates/sec

Title: US-09-733-306C-4
Perfect score: 1573
Sequence: 1 MANNDAVLKRLEQKGAEADQ.....FEVKGKGVCRAQTMSNSGIK 312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1573	100.0	312	5	ABB07442	Abb07442 Human p43
2	1573	100.0	312	6	AAO29575	Aao29575 Human Pc2
3	1573	100.0	312	8	ADH74487	Adh74487 Human wou
4	1573	100.0	312	8	ADP85653	Adp85653 Human end
5	1573	100.0	328	4	AAB63244	Aab63244 Human bre
6	1343	85.4	310	2	AAR72577	Aar72577 Mouse EMA
7	831	52.8	166	6	ABU10452	Abu10452 Immunolog
8	823	52.3	166	2	AAW14561	Aaw14561 Endotheli
9	709	45.1	147	6	ABU10250	Abu10250 Immunolog
10	671	42.7	294	4	ABB59210	Abb59210 Drosophil
11	522.5	33.2	280	7	ABM74361	Abm74361 DNA clone
12	522.5	33.2	289	7	ABM73655	Abm73655 DNA clone
13	516	32.8	108	6	ABU10451	Abu10451 Immunolog
14	510.5	32.5	659	5	AAE19811	Aae19811 Corn p012
15	507	32.2	95	8	ADN99968	Adn99968 Novel hum
16	499	31.7	797	8	ADN73157	Adn73157 Thale cre
17	451	28.7	536	4	AAB47611	Aab47611 Human ful
18	451	28.7	536	5	AAE13487	Aae13487 Human tyr
19	451	28.7	536	6	ABU72380	Abu72380 Human ful
20	450	28.6	528	7	ADF76697	Adf76697 Novel hum
21	450	28.6	528	8	ADN05830	Adn05830 Antipsori
22	449.5	28.6	527	7	ABU64299	Abu64299 Human tyr
23	440.5	28.0	301	2	AAW53962	Aaw53962 Monocyte
24	440.5	28.0	301	8	ADE76967	Ade76967 Human pro
25	430.5	27.4	178	4	AAB47613	Aab47613 Human Tyr
26	430.5	27.4	178	5	AAE13489	Aae13489 Human Tyr
27	430.5	27.4	178	6	ABU72382	Abu72382 Human Tyr
28	429.5	27.3	168	2	AAW06596	Aaw06596 Human end
29	429.5	27.3	168	3	AAY69674	Aay69674 Human end
30	400.5	25.5	382	3	AAG31613	Aag31613 Arabidops
31	400.5	25.5	389	3	AAG31612	Aag31612 Arabidops
32	400.5	25.5	440	3	AAG31611	Aag31611 Arabidops
33	395	25.1	525	4	ABB60745	Abb60745 Drosophil
34	391.5	24.9	265	7	ABM74280	Abm74280 DNA clone
35	343	21.8	273	3	AAG08528	Aag08528 Arabidops
36	294	18.7	123	8	ADN99487	Adn99487 Novel hum
37	293	18.6	58	6	ABU72393	Abu72393 Human pro
38	227	14.4	720	6	ADA33478	Ada33478 Acinetoba
39	225	14.3	690	6	ABU16644	Abu16644 Protein e
40	223.5	14.2	737	4	AAB96809	Aab96809 Putative
41	221.5	14.1	744	8	ADN46653	Adn46653 Thermococ
42	198.5	12.6	234	4	ABG22608	Abg22608 Novel hum
43	197.5	12.6	545	4	ABG17833	Abg17833 Novel hum
44	197	12.5	276	4	ABG26195	Abg26195 Novel hum
45	188.5	12.0	669	6	ABU29885	Abu29885 Protein e

ALIGNMENTS

RESULT 1
ABB07442
ID ABB07442 standard; protein; 312 AA.
XX
AC ABB07442;
XX
DT 23-APR-2002 (first entry)
XX
DE Human p43 polypeptide.
XX
KW Human; p43; cytostatic; anti-tumour; anti-angiogenic; EMAP II; cytokine;
KW chemotaxis.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Domain	1. .146
FT		/note= "N-terminal domain"
FT	Region	10. .21
FT		/note= "beta1 strand"
FT	Region	28. .34
FT		/note= "beta2 strand"
FT	Region	40. .46
FT		/note= "beta3 strand"
FT	Region	53. .56
FT		/note= "alpha 1 helix"
FT	Region	59. .66
FT		/note= "beta4 strand"
FT	Region	70. .72
FT		/note= "beta5i strand"
FT	Region	75. .77
FT		/note= "beta6i strand"
FT	Region	79. .85
FT		/note= "beta7 strand"
FT	Region	90. .92
FT		/note= "beta8 strand"
FT	Region	103. .106
FT		/note= "beta9 strand"
FT	Region	119. .123
FT		/note= "alpha 2 helix"
FT	Region	124. .130
FT		/note= "alpha 3 helix"
FT	Region	132. .134
FT		/note= "beta10 strand"
FT	Region	140. .142
FT		/note= "beta11 strand"
FT	Domain	147. .312
FT		/note= "C-terminal domain"

XX
PN WO200195927-A1.
XX
PD 20-DEC-2001.
XX
PF 14-JUN-2000; 2000WO-KR000630.
XX
PR 14-JUN-2000; 2000WO-KR000630.
XX
PA (IMAG-) IMAGENE CO LTD.
XX
PI Kim S, Ko Y, Kim YS, Jo YJ;
XX
DR WPI; 2002-098017/13.
DR N-PSDB; ABA94640.
XX
PT Novel anti-tumor and anti-angiogenic agent of p43 comprises N-terminal
PT domain and C-terminal domain containing eleven beta-strands forming a
PT structural core and three flanking alpha-helices.
XX
PS Claim 1; Fig 1; 35pp; English.
XX
CC The invention provides an anti-tumour and anti-angiogenic agent of p43
CC consisting of two domains, the N-terminal domain (146 amino acids) and C-
CC terminal domain (166 amino acids) containing 11 beta-strands forming a
CC structural core and 3 flanking alpha-helices. p43 is useful as an anti -
CC tumour and anti-angiogenic agent. p43 and its C-terminal cytokine domain
CC (EMAP II) induce regression of fibrosarcoma in immunocompromised mouse
CC while its N-terminal domain does not. p43 is a potent cytokine as
CC determined by the induction of tumour necrosis factor-alpha (TNF-alpha),
CC interleukin-6 (IL-6), IL-8 and matrix metalloprotease-9 or by its
CC activity of chemotaxis. The present sequence represents the human p43
CC polypeptide
XX
SQ Sequence 312 AA;

Query Match 100.0%; Score 1573; DB 5; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.5e-135;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

RESULT 2
AAO29575
ID AAO29575 standard; protein; 312 AA.
XX
AC AAO29575;
XX
DT 27-AUG-2003 (first entry)
XX
DE Human Pc240 protein.
XX
KW Human; differentially regulated protein; prevention; therapy; vaccine;
KW prostate cancer; endothelial monocyte activating polypeptide II;
KW gene therapy; Pc240.
XX
OS Homo sapiens.
XX
PN WO2003040331-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US035563.
XX
PR 07-NOV-2001; 2001US-0331041P.
PR 07-NOV-2001; 2001US-0331042P.
PR 18-DEC-2001; 2001US-0340251P.
PR 07-JAN-2002; 2002US-0344791P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX
DR WPI; 2003-449451/42.
DR N-PSDB; AAL60070.
XX
PT New polynucleotide for diagnosing, staging, monitoring, prognosticating,
PT preventing or treating, or determining the predisposition to, diseases or
PT conditions such as prostate cancer, and for research or forensic science.
XX
PS Disclosure; Page 183-184; 100pp; English.
XX
CC The present invention relates to novel differentially regulated genes and
CC polypeptides encoded by them. Sequences of the invention are useful in
CC diagnosing, staging, monitoring, prognosticating, preventing, treating or
CC determining the predisposition to diseases or conditions such as prostate
CC cancer. They may be used as molecular markers, drug targets, vaccines, in
CC gene therapy, research, clinical medicine or forensic science. The
CC present sequence is a differentially regulated prostate protein
CC (endothelial monocyte activating polypeptide II), Pc240
XX
SQ Sequence 312 AA;

Query Match 100.0%; Score 1573; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.5e-135;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

RESULT 3

ADH74487
ID ADH74487 standard; protein; 312 AA.
XX
AC ADH74487;
XX
DT 15-APR-2004 (first entry)
XX
DE Human wound healing stimulating polypeptide p43, SEQ ID 1.
XX
KW Vulnerary; antiulcer; dermatological; antiallergic; antiinflammatory;
KW wound healing; burn; ulcer; trauma; post-surgical; post-child birth;
KW chronic wound; dermatitis; sunburn; chemical burn; radiation burn;
KW thermal burn; pressure ulcer; plaster ulcer; decubitus ulcer; bedsores;
KW pressure sore; diabetes; poor circulation; impetigo; intertrigo;
KW folliculitis; eczema; p43.
XX
OS Homo sapiens.
XX
PN EP1384486-A1.
XX
PD 28-JAN-2004.
XX
PF 22-JUL-2003; 2003EP-00016724.
XX
PR 22-JUL-2002; 2002KR-00042858.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S;
XX
DR WPI; 2004-145613/15.
XX
PT Use of p43 polypeptide, for the manufacture of pharmaceutical composition
PT intended to stimulate wound healing in a subject.
XX
PS Claim 1; SEQ ID NO 1; 20pp; English.
XX
CC The invention relates to the use of one or more polypeptides (I) chosen
CC from a polypeptide having a fully defined p43 sequence of 312 amino acids
CC as given in the specification and a polypeptide having 70% or more
CC sequence homology with (S1), for the manufacture of the pharmaceutical
CC composition (II) which is intended to stimulate wound healing in a
CC subject. The polypeptide of the invention is useful for the manufacture
CC of the pharmaceutical composition (II) which is intended to stimulate
CC wound healing in a subject. The wound is chosen from burn, ulcer, trauma,
CC post-surgical, post-child birth, chronic wound and dermatitis. The burn
CC is chosen from sun burn, chemical burn, radiation burn, and thermal burn.
CC The ulcer is chosen from pressure ulcer, plaster ulcer and decubitus
CC ulcer. The chronic wound is chosen from bedsores, pressure sores,
CC diabetes-related and poor circulation-related. The dermatitis is chosen
CC from impetigo, intertrigo, folliculitis and eczema. The current sequence
CC represents the wound healing stimulating polypeptide of the invention.
XX
SQ Sequence 312 AA;

Query Match 100.0%; Score 1573; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.5e-135;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

RESULT 4
ADP85653
ID ADP85653 standard; protein; 312 AA.
XX
AC ADP85653;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human endothelial monocyte-activating polypeptide-II (EMAP-II).
XX
KW EMAP-II; endothelial monocyte-activating polypeptide-II; EMAP-2; SCYE1;
KW small inducible cytokine subfamily E member 1;
KW hyperproliferative disorder; cancer; gene therapy; human.
XX
OS Homo sapiens.
XX
PN US2004110144-A1.
XX
PD 10-JUN-2004.
XX
PF 09-DEC-2002; 2002US-00316232.
XX
PR 09-DEC-2002; 2002US-00316232.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Dean NM, Dobie KW;
XX
DR WPI; 2004-440333/41.
DR N-PSDB; ADP85584.
XX
PT New oligonucleotide compound that inhibits expression of EMAP-II, useful
PT for preparing a composition for treating hyperproliferative disorder,
PT e.g. cancer.
XX
PS Disclosure; Page 22-23; 35pp; English.
XX
CC The present invention relates to a compounds, compositions and methods
CC for modulating the expression of endothelial monocyte-activating
CC polypeptide-II (EMAP-II). EMAP-II is also known as EMAP-2, small
CC inducible cytokine subfamily E, member 1 (SCYE1). The compound comprises
CC antisense oligonucleotides targetted to EMAP-II. The invention is useful
CC for preparing a composition for treating hyperproliferative disorder e.g.
CC cancer. It is also useful in gene therapy. The present sequence is human
CC endothelial monocyte-activating polypeptide-II (EMAP-II). This sequence
CC is used in the invention.
XX
SQ Sequence 312 AA;

Query Match 100.0%; Score 1573; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.5e-135;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
                ||||||||||||
Db        301 CRAQTMSNSGIK 312


RESULT 5
AAB63244
ID   AAB63244 standard; protein; 328 AA.
XX
AC   AAB63244;
XX
DT   26-MAR-2001  (first entry)
XX
DE   Human breast cancer associated antigen protein sequence SEQ ID NO:606.
XX
KW   Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW   cancer associated antigen; cytostatic; cancer vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200073801-A2.
XX
PD   07-DEC-2000.
XX
PF   26-MAY-2000; 2000WO-US014749.
XX
PR   28-MAY-1999;   99US-0136526P.
PR   10-SEP-1999;   99US-0153454P.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Obata Y;
XX
DR   WPI; 2001-025274/03.
XX
PT   Nucleic acids encoding breast, gastric and prostate cancer associated
PT   antigen precursors, useful for diagnosing and treating a condition
PT   characterized by expression of an abnormal amount of a protein, e.g.
PT   cancer.
XX
PS   Example 1; Page 478; 799pp; English.
XX
CC   AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC   represent nucleotide sequences encoding human breast, gastric and
CC   prostate cancer associated antigen precursors (CAAP) respectively.
CC   AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC   represent human breast, gastric and prostate CAAP protein sequence
CC   respectively. CAAPs have cytostatic activity and can be used in the
CC   production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC   acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC   condition characterised by expression of an abnormal amount of a protein,
CC   e.g. cancer
XX
SQ   Sequence 328 AA;

  Query Match           100.0%;  Score 1573;  DB 4;  Length 328;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-135;
  Matches  312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 76

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 136

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 196

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 256

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 316

Qy        301 CRAQTMSNSGIK 312
                ||||||||||||
Db        317 CRAQTMSNSGIK 328


RESULT 6
AAR72577
ID   AAR72577 standard; protein; 310 AA.
XX
AC   AAR72577;
XX
DT   25-MAR-2003  (revised)
DT   29-SEP-1995  (first entry)
XX
DE   Mouse EMAPII.
XX
KW   EMAPII; endothelial monocyte activating polypeptide II; chemotaxis;
KW   inflammation; tissue factor; tumor; cancer; therapy; methA; sarcoma.
XX
OS   Mus musculus.
XX
PN   WO9509180-A1.
XX
PD   06-APR-1995.
XX
PF   29-SEP-1994;   94WO-US011085.
XX
PR   29-SEP-1993;   93US-00129456.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Stern DM,  Clauss M,  Kao J,  Kayton M,  Libutti SK;
XX
DR   WPI; 1995-147389/19.
DR   N-PSDB; AAQ86718.
XX
PT   New endothelial monocyte activating polypeptide II - induces chemotaxis,
PT   inflammation and tissue factor, useful for treating tumours, also related
PT   antibodies, DNA and active fragments.
XX
PS   Disclosure; Fig 4; 180pp; English.
XX
CC   A mouse methA sarcoma cDNA library was screened with a probe based on the
CC   N-terminal sequence of mouse EMAPII. Overlapping clones were combined to
CC   obtain a contigous full-length sequence (given in AAQ86718) encoding a 33
CC   kDa protein (AAR72577). Recombinant EMAPII was expressed in E. coli.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 310 AA;

  Query Match           85.4%;  Score 1343;  DB 2;  Length 310;
  Best Local Similarity  85.9%;  Pred. No. 1.9e-114;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
                || ||||||||||||||||||||||||||:||||||||||:||||||||||||||||||||
```

```
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |    ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||:|||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310
```

RESULT 7
ABU10452
ID ABU10452 standard; protein; 166 AA.
XX
AC ABU10452;
XX
DT 01-AUG-2003 (first entry)
XX
DE Immunological enhancement agent p43 (91-256).
XX
KW Immunostimulant; cytokine agonist; p43; cytokine; interleukin-8; TNF;
KW tumour necrosis factor.
XX
OS Mammalia.
XX
PN US2003004309-A1.
XX
PD 02-JAN-2003.
XX
PF 16-AUG-2001; 2001US-00930169.
XX
PR 05-JUN-2001; 2001KR-00031310.
XX
PA (KIMS/) KIM S.
PA (KOYY/) KO Y.
XX
PI Kim S, Ko Y;
XX
DR WPI; 2003-447359/42.
XX
PT New immunological enhancement agent comprising an N-terminal peptide of
PT p43, useful for increasing the amount of tumor necrosis factor and
PT interleukin-8 to improve the immune response.
XX
PS Claim 3; Page 5; 12pp; English.
XX
CC The present sequence represents an immunological enhancement agent
CC comprising an N-terminal peptide of p43. The peptide can act as a
CC cytokine to increase the amount of tumour necrosis factor (TNF) and
CC interleukin-8 useful for improving an immune response and used as an
CC immunological enhancement agent. The present sequence represents the
CC amino acid sequence of the immunological enhancement agent p43 (91-256)
XX
SQ Sequence 166 AA;

Query Match 52.8%; Score 831; DB 6; Length 166;
Best Local Similarity 100.0%; Pred. No. 7.4e-68;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         91 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 60

Qy        151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 120

Qy        211 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 256
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 166
```

RESULT 8
AAW14561
ID AAW14561 standard; protein; 166 AA.
XX
AC AAW14561;
XX
DT 01-DEC-1997 (first entry)
XX
DE Endothelial monocyte activating polypeptide II.
XX
KW EMAP-II; inhibitor; endothelial cell formation; blood vessel;
KW retinopathy; tumour; subcutaneous; intraperitoneal; intravenous;
KW intracranial; intratumoural; carcinoma; glioblastoma;
KW positive pressure microinfusion; aortic cell; angiogenesis; eye disease;
KW diabetes; sickle cell anaemia; prematurity;
KW age-related macular degeneration.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /label= Ser, Met, Gly
FT Misc-difference 15
FT /label= Cys, Arg
XX
PN WO9710841-A1.
XX
PD 27-MAR-1997.
XX
PF 18-SEP-1996; 96WO-US015007.
XX
PR 18-SEP-1995; 95US-0003898P.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Stern D, Schwarz M;
XX
DR WPI; 1997-202617/18.
XX
PT Treatment of tumours with endothelial monocyte activating peptide II -
PT also used to inhibit growth of endothelial cells and formation of blood
PT vessels, e.g. in retinopathy.
XX
PS Claim 13; Page 73; 104pp; English.
XX
CC A method has been developed for the treatment of tumours. The method
CC involves subcutaneous, intraperitoneal, intravenous, intracranial or
CC intratumoural administration of an endothelial monocyte activating
CC polypeptide II (EMAP-II), or a EMAP-II-derived polypeptide. The present
CC sequence represents murine EMAP-II which can be used to produce the EMAP-
CC II-derived polypeptide as long as it contains at least ninety percent
CC homology to the present sequence. The method is especially used to treat
CC carcinoma (administered intraperitoneally) or glioblastoma (administered
CC intratumourally by positive pressure microinfusion). Inhibition of
CC endothelial cell growth is particularly applied to aortic cells and
CC inhibition of angiogenesis to treatment of eye diseases associated with
CC excessive blood vessel formation, especially retinopathy (diabetic,
CC related to sickle cell anaemia, retinopathy of prematurity or age-related
CC macular degeneration). Intraperitoneal administration allows tumours too
CC small (not over 2 mm in diameter) for intratumoural injection to be
CC treated before they grow larger

```
XX
SQ   Sequence 166 AA;

  Query Match             52.3%;  Score 823;  DB 2;  Length 166;
  Best Local Similarity   93.9%;  Pred. No. 4e-67;
  Matches  155;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        148 KPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMV 207
              |||| |||||||| |:||:|||||||||||||||| ||||||||||||||||||||||
Db          2 KPIDASRLDLRIGXIVTAKKHPDADSLYVEEVDVGEAAPRTVVSGLVNHVPLEQMQNRMV 61

Qy        208 ILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNPK 267
              :|||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db         62 VLLCNLKPAKMRGVLSQAMVMCASSPEKVEILAPPNGSVPGDRITFDAFPGEPDKELNPK 121

Qy        268 KKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              |||||||||||||| |||||||| |||||||||||||:|||||
Db        122 KKIWEQIQPDLHTNAECVATYKGAPFEVKGKGVCRAQTMANSGIK 166


RESULT 9
ABU10250
ID   ABU10250 standard; protein; 147 AA.
XX
AC   ABU10250;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Immunological enhancement agent p43 (1-147).
XX
KW   Immunostimulant; cytokine agonist; p43; cytokine; interleukin-8; TNF;
KW   tumour necrosis factor.
XX
OS   Mammalia.
XX
PN   US2003004309-A1.
XX
PD   02-JAN-2003.
XX
PF   16-AUG-2001; 2001US-00930169.
XX
PR   05-JUN-2001; 2001KR-00031310.
XX
PA   (KIMS/) KIM S.
PA   (KOYY/) KO Y.
XX
PI   Kim S,  Ko Y;
XX
DR   WPI; 2003-447359/42.
XX
PT   New immunological enhancement agent comprising an N-terminal peptide of
PT   p43, useful for increasing the amount of tumor necrosis factor and
PT   interleukin-8 to improve the immune response.
XX
PS   Claim 1; Page 4; 12pp; English.
XX
CC   The present sequence represents an immunological enhancement agent
CC   comprising an N-terminal peptide of p43. The peptide can act as a
CC   cytokine to increase the amount of tumour necrosis factor (TNF) and
CC   interleukin-8 useful for improving an immune response and used as an
CC   immunological enhancement agent. The present sequence represents the
CC   amino acid sequence of the immunological enhancement agent p43 (1-147)
XX
SQ   Sequence 147 AA;

  Query Match             45.1%;  Score 709;  DB 6;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-57;
  Matches  147;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147
              |||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147


RESULT 10
ABB59210
ID   ABB59210 standard; protein; 294 AA.
XX
AC   ABB59210;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 4422.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL03313.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 4422; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 294 AA;

  Query Match             42.7%;  Score 671;  DB 4;  Length 294;
  Best Local Similarity   44.9%;  Pred. No. 8.2e-53;
  Matches  141;  Conservative   57;  Mismatches   88;  Indels   28;  Gaps    5;

Qy          6 AVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIEELKQE 65
              | |:::      |: :|  :: ::| :::    |   |::::|  ||| | ||:|    :
Db          2 ADLQQIASNNERAEALINSIEAEISGIQQ----QLVERQKQELIKENAALAKEVEAALAQ 57

Qy         66 LIQAEIQNGVKQIAFP-------SGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTG 118
              |:| |::|| ||| |        |  |:    ::      :||     :
Db         58 LVQLELRNGKKQIPVPGARGFCTSAAPV----VMPAEAGPATAAPAAPA---------- 102

Qy        119 DEKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEE 178
                | |||  ||| ::|:   :  | :|   |:|| |||||:| |:   :|||||||:|:
```

```
Db        103 -PKPAKEPKEKKSKEKKPAAEKPA-AAPEAPVDVGRLDLRVGKIVEVGRHPDADSLYLEK 160

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEI 238
              :| || |||||||||| ||||:||||:|:::|||||||||||| |:|||||||:|||:|:
Db        161 IDCGEAAPRTVVSGLVKFVPLEEMQNRLVVVMCNLKPAKMRGVTSEAMVMCASTPEKVEV 220

Qy        239 LAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGK 298
              |:|| |:|||| :  : :| :|| :|||||||:|   ||| || | || |||    | ||
Db        221 LSPPAGAVPGDLVHCEGYPRQPDAQLNPKKKIFESCAPDLKTNGELVACYKGAALHVPGK 280

Qy        299 GVCRAQTMSNSGIK 312
              |   |||: |  :|
Db        281 GNVVAQTLKNVNVK 294


RESULT 11
ABM74361
ID   ABM74361 standard; protein; 280 AA.
XX
AC   ABM74361;
XX
DT   17-OCT-2003  (first entry)
XX
DE   DNA clone originating in barley containing SNP sequence #771.
XX
KW   Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX
OS   Hordeum vulgare.
XX
PN   WO2003057877-A1.
XX
PD   17-JUL-2003.
XX
PF   16-DEC-2002; 2002WO-IB005403.
XX
PR   20-DEC-2001; 2001JP-00387059.
PR   20-DEC-2001; 2001JP-00387131.
PR   20-DEC-2001; 2001JP-00403299.
PR   20-DEC-2001; 2001JP-00403300.
PR   27-SEP-2002; 2002JP-00327515.
XX
PA   (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI   Sato K,  Takeda K,  Kohara Y;
XX
DR   WPI; 2003-587127/55.
XX
PT   Single nucleotide polymorphism sites in barley varieties and DNA
PT   sequences containing them for analysis and identification of barley
PT   varieties and production of barley transformants with desired
PT   characteristics.
XX
PS   Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC   The present invention relates to oligonucleotide clones originating in
CC   barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC   (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC   varieties, identification of particular varieties and genotype-phenotype
CC   analysis, isolation of specific genes and creation of new varieties by
CC   transformation of barley varieties with them and production of new barley
CC   varieties with desired properties. The present sequence represents an
CC   oligonucleotide clone sequence featured in the specification. The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published-pct-sequences
XX
SQ   Sequence 280 AA;

  Query Match           33.2%;  Score 522.5;  DB 7;  Length 280;
  Best Local Similarity   45.4%;  Pred. No. 3.2e-39;
  Matches  108;  Conservative   42;  Mismatches   77;  Indels   11;  Gaps    5;

Qy         83 GTPLHANSMVSENVIQSTAVTTVSSGTKEQIKG-GTGDEKKAKEKIEKKGEK-----KEK 136
               | |:     : :  :::  :    |  :  :||    : ||  ||:  || |      :|
Db         46 GKPVPLFKELKDQEVEAFRIKFAGSQAERILKGQADAEAKKVAEKL--KGTKLSEGSSKK 103

Qy        137 KQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNH 196
              ||   : |  |: : |::||:|:| |  | ||||||||||||:|| | ||||||||||
Db        104 KQSGGSKSKTSEDVSVAKLDIRVGLIRKAEKHPDADSLYVEEIDVVEEAPRTVVSGLVKF 163

Qy        197 VPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAPPNGSVPGDRITFD 254
              :|||:|||| | :|||||  |||: | |||: ||: :  |:|:: ||  :  |:|:||
Db        164 IPLEEMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDHTKVELVEPPESAAVGERVTFA 223

Qy        255 AFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              | |||:  || | | ||::  |||:| | || || |||     |||: :|::|  |:
Db        224 GFSGEPEASLNAKSKTWEKLSADLHSNGELVACYKDVPFTTSA-GVCKVKTIANGAIR 280


RESULT 12
ABM73655
ID   ABM73655 standard; protein; 289 AA.
XX
AC   ABM73655;
XX
DT   17-OCT-2003  (first entry)
XX
DE   DNA clone originating in barley containing SNP sequence #65.
XX
KW   Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX
OS   Hordeum vulgare.
XX
PN   WO2003057877-A1.
XX
PD   17-JUL-2003.
XX
PF   16-DEC-2002; 2002WO-IB005403.
XX
PR   20-DEC-2001; 2001JP-00387059.
PR   20-DEC-2001; 2001JP-00387131.
PR   20-DEC-2001; 2001JP-00403299.
PR   20-DEC-2001; 2001JP-00403300.
PR   27-SEP-2002; 2002JP-00327515.
XX
PA   (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI   Sato K,  Takeda K,  Kohara Y;
XX
DR   WPI; 2003-587127/55.
XX
PT   Single nucleotide polymorphism sites in barley varieties and DNA
PT   sequences containing them for analysis and identification of barley
PT   varieties and production of barley transformants with desired
PT   characteristics.
XX
PS   Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC   The present invention relates to oligonucleotide clones originating in
CC   barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC   (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC   varieties, identification of particular varieties and genotype-phenotype
CC   analysis, isolation of specific genes and creation of new varieties by
CC   transformation of barley varieties with them and production of new barley
CC   varieties with desired properties. The present sequence represents an
CC   oligonucleotide clone sequence featured in the specification. The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published-pct-sequences
XX
SQ   Sequence 289 AA;

  Query Match           33.2%;  Score 522.5;  DB 7;  Length 289;
  Best Local Similarity   45.4%;  Pred. No. 3.4e-39;
  Matches  108;  Conservative   42;  Mismatches   77;  Indels   11;  Gaps    5;
```

```
Qy          83 GTPLHANSMVSENVIQSTAVTTVSSGTKEQIKG-GTGDEKKAKEKIEKKGEK-----KEK 136
               | |:     : :  :::  :    |  :  :||    : ||  ||:  || |      :|
Db          55 GKPVPLFKELKDQEVEAFRIKFAGSQAERILKGQADAEAKKVAEKL--KGTKLSEGSSKK 112

Qy         137 KQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNH 196
               ||   : |  |: : |::||:|:| |  | ||||||||||||:|| | ||||||||||
Db         113 KQSGGSKSKTSEDVSVAKLDIRVGLIRKAEKHPDADSLYVEEIDVVEEAPRTVVSGLVKF 172

Qy         197 VPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAPPNGSVPGDRITFD 254
               :|||:|||| | :|||||  |||: | |||: ||: :  |:|:: ||  :  |:|:||
Db         173 IPLEEMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDHTKVELVEPPESAAVGERVTFA 232

Qy         255 AFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
               | |||:  || | | ||::  |||:| | || || |||     |||: :|::|  |:
Db         233 GFSGEPEASLNAKSKTWEKLSADLHSNGELVACYKDVPFTTSA-GVCKVKTIANGAIR 289
```

RESULT 13
ABU10451
ID ABU10451 standard; protein; 108 AA.
XX
AC ABU10451;
XX
DT 01-AUG-2003 (first entry)
XX
DE Immunological enhancement agent p43 (1-108).
XX
KW Immunostimulant; cytokine agonist; p43; cytokine; interleukin-8; TNF;
KW tumour necrosis factor.
XX
OS Mammalia.
XX
PN US2003004309-A1.
XX
PD 02-JAN-2003.
XX
PF 16-AUG-2001; 2001US-00930169.
XX
PR 05-JUN-2001; 2001KR-00031310.
XX
PA (KIMS/) KIM S.
PA (KOYY/) KO Y.
XX
PI Kim S, Ko Y;
XX
DR WPI; 2003-447359/42.
XX
PT New immunological enhancement agent comprising an N-terminal peptide of
PT p43, useful for increasing the amount of tumor necrosis factor and
PT interleukin-8 to improve the immune response.
XX
PS Claim 2; Page 5; 12pp; English.
XX
CC The present sequence represents an immunological enhancement agent
CC comprising an N-terminal peptide of p43. The peptide can act as a
CC cytokine to increase the amount of tumour necrosis factor (TNF) and
CC interleukin-8 useful for improving an immune response and used as an
CC immunological enhancement agent. The present sequence represents the
CC amino acid sequence of the immunological enhancement agent p43 (1-108)
XX
SQ Sequence 108 AA;

Query Match 32.8%; Score 516; DB 6; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.2e-39;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy          61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108
```

RESULT 14
AAE19811
ID AAE19811 standard; protein; 659 AA.
XX
AC AAE19811;
XX
DT 18-JUN-2002 (first entry)
XX
DE Corn p0128.cpibk48r clone methionyl-tRNA synthetase (MTS).
XX
KW Corn; methionine metabolic enzyme; methionyl-tRNA synthetase; MTS;
KW methionyl-tRNA formyltransferase; MTF; transgenic plant; gene mapping;
KW plant breeding; allele specific ligation; radiation hybrid mapping;
KW nucleotide extension reaction; allele-specific amplification; enzyme;
KW physical mapping; protein; EC 6.1.1.10.
XX
OS Zea mays.
XX
PN US6346403-B1.
XX
PD 12-FEB-2002.
XX
PF 07-SEP-1999; 99US-00392772.
XX
PR 08-SEP-1998; 98US-0099519P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Rafalski JA, Famodu LO;
XX
DR WPI; 2002-224991/28.
DR N-PSDB; AAD31528.
XX
PT New isolated polynucleotide encoding plant methionyl-tRNA
PT formyltransferase or methionyl-tRNA synthetase polypeptide, useful for
PT producing transgenic plants and as probes or primers.
XX
PS Claim 1; Col 37-40; 24pp; English.
XX
CC The invention relates to nucleic acids encoding methionine metabolic
CC enzymes, methionyl-tRNA formyltransferase (MTF) and methionyl-tRNA
CC synthetase (MTS). The nucleic acids of the invention are useful for
CC selecting an isolated polynucleotide that affects the level of expression
CC of MTS or MTF polypeptide in a cell. They are useful as a probe or primer
CC for obtaining a nucleic acid fragment encoding a substantial portion of
CC MTF or MTS, preferably plant MTF or plant MTS gene. They are also useful
CC for creating transgenic plants in which the MTS or MTF polypeptides are
CC present at higher or lower levels than normal or in cell types or in
CC developmental stages in which they are not normally found. These
CC sequences may also be used as probes for genetically and physically
CC mapping the genes that they are a part of and used as markers for traits
CC linked to those genes. This information is used in plant breeding in
CC order to develop lines with desired phenotypes. These sequences are also
CC used in a variety of amplification based methods of genetic and physical
CC mapping such as allele-specific amplification, polymorphism of PCR
CC amplified fragments, allele specific ligation, nucleotide extension
CC reactions and radiation hybrid mapping. The present sequence is corn MTS
CC enzyme
XX
SQ Sequence 659 AA;

Query Match 32.5%; Score 510.5; DB 5; Length 659;
Best Local Similarity 38.6%; Pred. No. 1.4e-37;
Matches 123; Conservative 49; Mismatches 84; Indels 63; Gaps 11;

```
Qy          18 ADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKE----------------IEE 61
               :|:::   :|::|  |: :  :| |||    | || |                  :|
Db         380 SDEVL----RQLNLSPEENL---SLSEEKG---EIAKAKTPWNFVPAGHRIGKPAPLFKE 429

Qy          62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGD 119
               || |                 ||         :|:     :   |   |:||
```

```
Db        430 LKDE-----------------DVALHREKYAGSQAERSSKAAADAEANKVANQLKG---- 468

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLY 175
                  |:   | ||| |:||  | : ||     | |::||:|:| |  | |||||||||
Db        469 -----TKLSDGGTKKEPKKQS--GGSKSKTAEADITVAKLDIRVGLIRKAEKHPDADSLY 521

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:|||| ||||||||||  :|||:||:| | :|||||  |||: | |||: ||: :
Db        522 VEEIDVGEDAPRTVVSGLVKFIPLEEMQSRKVCVLCNLKPVAMRGIKSHAMVLAASNEDH 581

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
               |:|:: ||  :  |:|:||  : |||:  |: | | ||::  :||:| | || |: |||
Db        582 TKVELVEPPESAAVGERVTFAGYAGEPEASLSGKSKTWEKLAAELHSNGELVACYRDVPF 641

Qy        294 EVKGKGVCRAQTMSNSGIK 312
                  |||| :|::|  |:
Db        642 TTSA-GVCRVKTIANGEIR 659
```

RESULT 15
ADN99968
ID ADN99968 standard; protein; 95 AA.
XX
AC ADN99968;
XX
DT 29-JUL-2004 (first entry)
XX
DE Novel human protein sequence #784.
XX
KW anti-inflammatory; dermatological; neuroprotective; immunomodulator;
KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;
KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
KW early aging; hormonal imbalance; ischemic heart disease;
KW ulcerative colitis.
XX
OS Homo sapiens.
XX
PN WO2004038003-A2.
XX
PD 06-MAY-2004.
XX
PF 24-OCT-2003; 2003WO-US033947.
XX
PR 25-OCT-2002; 2002US-0421061P.
PR 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 25-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426355P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429302P.
PR 27-NOV-2002; 2002US-0429326P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
DR WPI; 2004-365511/34.
DR N-PSDB; ADN99184.
XX
PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
PS Claim 14; SEQ ID NO 1568; 532pp; English.
XX
CC The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
SQ Sequence 95 AA;

```
  Query Match             32.2%;  Score 507;  DB 8;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-38;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        218 MRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPD 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPD 60

Qy        278 LHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              |||||||||||||||||||||||||||||||||||
Db         61 LHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 95
```

Search completed: January 14, 2005, 15:55:12
Job time : 163 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 15:49:58 ; Search time 40 Seconds
(without alignments)
517.280 Million cell updates/sec

Title: US-09-733-306C-4
Perfect score: 1573
Sequence: 1 MANNDAVLKRLEQKGAEADQ.....FEVKGKGVCRAQTMSNSGIK 312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1568	99.7	312	3	US-08-360-821B-36	Sequence 36, Appl
2	1568	99.7	312	4	US-09-851-026-36	Sequence 36, Appl
3	1510	96.0	310	1	US-08-129-456A-37	Sequence 37, Appl
4	1343	85.4	310	1	US-08-129-456A-36	Sequence 36, Appl
5	1343	85.4	310	2	US-08-705-868-3	Sequence 3, Appli
6	1343	85.4	310	3	US-09-123-615-3	Sequence 3, Appli
7	1343	85.4	310	3	US-08-360-821B-35	Sequence 35, Appl
8	1343	85.4	310	4	US-09-851-026-35	Sequence 35, Appl
9	931	59.2	183	3	US-08-483-534A-7	Sequence 7, Appli
10	835	53.1	166	1	US-08-129-456A-11	Sequence 11, Appl
11	835	53.1	166	3	US-08-360-821B-11	Sequence 11, Appl
12	835	53.1	166	4	US-09-851-026-11	Sequence 11, Appl
13	510.5	32.5	659	3	US-09-392-772-10	Sequence 10, Appl
14	440.5	28.0	301	2	US-08-705-868-1	Sequence 1, Appli
15	440.5	28.0	301	3	US-09-123-615-1	Sequence 1, Appli
16	440.5	28.0	301	4	US-09-919-039-132	Sequence 132, App
17	438.5	27.9	215	4	US-09-248-796A-19812	Sequence 19812, A
18	429.5	27.3	168	3	US-08-483-534A-2	Sequence 2, Appli
19	235	14.9	174	4	US-09-248-796A-19813	Sequence 19813, A
20	227	14.4	720	4	US-09-328-352-4765	Sequence 4765, Ap
21	188.5	12.0	674	4	US-09-107-532A-6201	Sequence 6201, Ap
22	186.5	11.9	648	1	US-08-451-715A-4	Sequence 4, Appli
23	182.5	11.6	111	4	US-09-489-039A-9293	Sequence 9293, Ap
24	174.5	11.1	493	4	US-09-134-000C-4034	Sequence 4034, Ap
25	157	10.0	665	3	US-08-844-059-2	Sequence 2, Appli
26	157	10.0	665	3	US-09-431-202-2	Sequence 2, Appli
27	152.5	9.7	679	4	US-09-583-110-2751	Sequence 2751, Ap
28	147	9.3	221	4	US-09-252-991A-21654	Sequence 21654, A
29	136	8.6	716	4	US-09-489-039A-8243	Sequence 8243, Ap
30	134.5	8.6	680	4	US-09-252-991A-17566	Sequence 17566, A
31	129	8.2	690	4	US-09-540-236-3507	Sequence 3507, Ap
32	128	8.1	115	3	US-08-844-059-4	Sequence 4, Appli
33	128	8.1	115	3	US-09-431-202-4	Sequence 4, Appli
34	128	8.1	683	4	US-09-543-681A-5403	Sequence 5403, Ap
35	119.5	7.6	589	4	US-09-543-681A-7643	Sequence 7643, Ap
36	117	7.4	721	3	US-09-134-078-19	Sequence 19, Appl
37	116.5	7.4	414	5	PCT-US93-03077-3	Sequence 3, Appli
38	116.5	7.4	1093	5	PCT-US93-03077-1	Sequence 1, Appli
39	115.5	7.3	1493	4	US-09-713-273A-20	Sequence 20, Appl
40	115	7.3	890	4	US-09-849-602-19	Sequence 19, Appl
41	114.5	7.3	271	1	US-08-664-596B-11	Sequence 11, Appl
42	114.5	7.3	431	4	US-09-286-981B-3	Sequence 3, Appli
43	113.5	7.2	618	4	US-09-710-279-860	Sequence 860, App
44	113.5	7.2	660	3	US-09-134-001C-3350	Sequence 3350, Ap
45	113	7.2	127	4	US-09-543-681A-7844	Sequence 7844, Ap

ALIGNMENTS

RESULT 1
US-08-360-821B-36
; Sequence 36, Application US/08360821B
; Patent No. 6228837
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Clauss, Matthias
; APPLICANT: Kao, Janet
; APPLICANT: Kayton, Mark
; APPLICANT: Libutti, Steven K
; TITLE OF INVENTION: Endothelial Monocyte Activating
; TITLE OF INVENTION: Polypeptide II: A Mediator Which Activates Host Response
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,821B
; FILING DATE: 08-OCT-96
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-360-821B-36

Query Match 99.7%; Score 1568; DB 3; Length 312;
Best Local Similarity 99.7%; Pred. No. 2.7e-136;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
                 |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIPFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
                 ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

RESULT 2
US-09-851-026-36
; Sequence 36, Application US/09851026
; Patent No. 6734168
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; Clauss, Matthias
; Kao, Janet
; Kayton, Mark
; Libutti, Steven K
; TITLE OF INVENTION: Endothelial Monocyte Activating
; Polypeptide II: A Mediator Which Activates Host Response
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,026
; FILING DATE: 07-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,821
; FILING DATE: 08-OCT-96
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-851-026-36

```
Query Match           99.7%;  Score 1568;  DB 4;  Length 312;
Best Local Similarity  99.7%;  Pred. No. 2.7e-136;
Matches  311;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
                 |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIPFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
                 ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

RESULT 3
US-08-129-456A-37
; Sequence 37, Application US/08129456A
; Patent No. 5641867
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; APPLICANT: Clauss, Matthias
; APPLICANT: Kao, Janet
; APPLICANT: Kayton, Mark
; APPLICANT: Libutti, Steven K.
; TITLE OF INVENTION: Endothelial-Monocyte Activating
; TITLE OF INVENTION: Polypeptide II: A Mediator Which
; TITLE OF INVENTION: Activates Host Response
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0 Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/129,456A
; FILING DATE: 29-SEP-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 278 0400
; TELEFAX: 212 391 0526
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-129-456A-37

```
  Query Match           96.0%;  Score 1510;  DB 1;  Length 310;
  Best Local Similarity  97.8%;  Pred. No. 5.9e-131;
  Matches  306;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||   ||
Db         61 ELKQELIQAEIQNGVKQIPFPSGTPLHANSMVSENVIQSTAVATTVSSGTKEQIK--AGD 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        119 EKKAKEKIEKKGEKK-KKQQSIAGSADSKPIDASRLDLRIGCIITARKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||||||||
Db        298 VCRAQTMSNSGIK 310


RESULT 4
US-08-129-456A-36
; Sequence 36, Application US/08129456A
; Patent No. 5641867
;  GENERAL INFORMATION:
;    APPLICANT:  Stern, David M.
;    APPLICANT:  Clauss, Matthias
;    APPLICANT:  Kao, Janet
;    APPLICANT:  Kayton, Mark
;    APPLICANT:  Libutti, Steven K.
;    TITLE OF INVENTION:  Endothelial-Monocyte Activating
;    TITLE OF INVENTION:  Polypeptide II: A Mediator Which
;    TITLE OF INVENTION:  Activates Host Response
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0 Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/129,456A
;      FILING DATE:  29-SEP-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  41735
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 278 0400
;      TELEFAX:  212 391 0526
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  310 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein

US-08-129-456A-36

  Query Match           85.4%;  Score 1343;  DB 1;  Length 310;
  Best Local Similarity  85.9%;  Pred. No. 1.4e-115;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||:||||||||||:||||||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||||:|||||||||||||||||||||||||||:|||||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 5
US-08-705-868-3
; Sequence 3, Application US/08705868
; Patent No. 5885798
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL MONOCYTE ACTIVATING CYTOKINE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/705,868
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0117 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  310 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
```

```
;      CLONE:  498910
US-08-705-868-3

  Query Match             85.4%;  Score 1343;  DB 2;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-115;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| ||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 6
US-09-123-615-3
; Sequence 3, Application US/09123615
; Patent No. 6090377
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL MONOCYTE ACTIVATING CYTOKINE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/123,615
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,868
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0117 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  310 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  498910
US-09-123-615-3

  Query Match             85.4%;  Score 1343;  DB 3;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-115;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| ||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              |||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 7
US-08-360-821B-35
; Sequence 35, Application US/08360821B
; Patent No. 6228837
;  GENERAL INFORMATION:
;    APPLICANT:  Stern, David M.
;    APPLICANT:  Clauss, Matthias
;    APPLICANT:  Kao, Janet
;    APPLICANT:  Kayton, Mark
;    APPLICANT:  Libutti, Steven K.
;    TITLE OF INVENTION:  Endothelial Monocyte Activating
;    TITLE OF INVENTION:  Polypeptide II:  A Mediator Which Activates Host Response
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham, LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.30, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/360,821B
;      FILING DATE:  08-OCT-96
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  41735-A-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  35:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  310 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-360-821B-35

  Query Match             85.4%;  Score 1343;  DB 3;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-115;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 8
US-09-851-026-35
; Sequence 35, Application US/09851026
; Patent No. 6734168
;   GENERAL INFORMATION:
;        APPLICANT: Stern, David M.
;                   Clauss, Matthias
;                   Kao, Janet
;                   Kayton, Mark
;                   Libutti, Steven K
;        TITLE OF INVENTION: Endothelial Monocyte Activating
;                            Polypeptide II:  A Mediator Which Activates Host Response
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Cooper & Dunham, LLP
;             STREET: 1185 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.30, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/851,026
;             FILING DATE: 07-May-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/360,821
;             FILING DATE: 08-OCT-96
;        ATTORNEY/AGENT INFORMATION:
;             NAME: White, John P.
;             REGISTRATION NUMBER: 28,678
;             REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-278-0400
;             TELEFAX: 212-391-0525
;   INFORMATION FOR SEQ ID NO: 35:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 310 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-851-026-35

  Query Match             85.4%;  Score 1343;  DB 4;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 1.4e-115;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 9
US-08-483-534A-7
; Sequence 7, Application US/08483534A
; Patent No. 6013483
;  GENERAL INFORMATION:
;    APPLICANT:  Coleman, Timothy A
;    APPLICANT:  Rosen, Craig
;    TITLE OF INVENTION:  Endothelial-Monocyte Activating
;    TITLE OF INVENTION:  Polypeptide III
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,534A
;      FILING DATE:  07 JUN 95
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-464 (PF206)
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  183 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-483-534A-7

  Query Match             59.2%;  Score 931;  DB 3;  Length 183;
  Best Local Similarity   98.4%;  Pred. No. 6e-78;
  Matches  180;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        130 KGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTV 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 KGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTV 60

Qy        190 VSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGD 249
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVKSQAMVMCASSPEKIEILAPPNGSVPGD 120

Qy        250 RITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNS 309
              ||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db        121 RITFDAFPGEPDKEKNPKKKIWEQIQPDLHTNDECVATYKEVPFEVKGKGVCRAQTMSNS 180

Qy        310 GIK 312
              |||
Db        181 GIK 183


RESULT 10
US-08-129-456A-11
; Sequence 11, Application US/08129456A
; Patent No. 5641867
;  GENERAL INFORMATION:
;    APPLICANT:  Stern, David M.
;    APPLICANT:  Clauss, Matthias
;    APPLICANT:  Kao, Janet
;    APPLICANT:  Kayton, Mark
;    APPLICANT:  Libutti, Steven K.
;    TITLE OF INVENTION:  Endothelial-Monocyte Activating
;    TITLE OF INVENTION:  Polypeptide II: A Mediator Which
;    TITLE OF INVENTION:  Activates Host Response
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0 Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/129,456A
;      FILING DATE:  29-SEP-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  41735
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 278 0400
;      TELEFAX:  212 391 0526
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Peptide
US-08-129-456A-11

  Query Match             53.1%;  Score 835;  DB 1;  Length 166;
  Best Local Similarity   94.0%;  Pred. No. 3.6e-69;
  Matches  156;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        147 SKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRM 206
              ||||| |||||||||||:||:||||||||||||||||| ||||||||||||||||||||||
Db          1 SKPIDASRLDLRIGCIVTAKKHPDADSLYVEEVDVGEAAPRTVVSGLVNHVPLEQMQNRM 60

Qy        207 VILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNP 266
              |:||||||||||||||||||||||||:|:|||||||||||||||||||||||||||||||
Db         61 VVLLCNLKPAKMRGVLSQAMVMCASSPDKVEILAPPNGSVPGDRITFDAFPGEPDKELNP 120

Qy        267 KKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              ||||||||||||||| |||||||| |||||||||||||||:|||||
Db        121 KKKIWEQIQPDLHTNAECVATYKGAPFEVKGKGVCRAQTMANSGIK 166


RESULT 11
US-08-360-821B-11
; Sequence 11, Application US/08360821B
; Patent No. 6228837
;  GENERAL INFORMATION:
;    APPLICANT:  Stern, David M.
;    APPLICANT:  Clauss, Matthias
;    APPLICANT:  Kao, Janet
;    APPLICANT:  Kayton, Mark
;    APPLICANT:  Libutti, Steven K
;    TITLE OF INVENTION:  Endothelial Monocyte Activating
;    TITLE OF INVENTION:  Polypeptide II:  A Mediator Which Activates Host Response
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham, LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.30, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/360,821B
;      FILING DATE:  08-OCT-96
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  41735-A-PCT-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0525
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Peptide
US-08-360-821B-11

  Query Match             53.1%;  Score 835;  DB 3;  Length 166;
  Best Local Similarity   94.0%;  Pred. No. 3.6e-69;
  Matches  156;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        147 SKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRM 206
              ||||| |||||||||||:||:||||||||||||||||| ||||||||||||||||||||||
```

```
Db          1 SKPIDASRLDLRIGCIVTAKKHPDADSLYVEEVDVGEAAPRTVVSGLVNHVPLEQMQNRM 60

Qy        207 VILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNP 266
              |:|||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||
Db         61 VVLLCNLKPAKMRGVLSQAMVMCASSPDKVEILAPPNGSVPGDRITFDAFPGEPDKELNP 120

Qy        267 KKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              ||||||||||||||| |||||||| ||||||||||||||:|||||
Db        121 KKKIWEQIQPDLHTNAECVATYKGAPFEVKGKGVCRAQTMANSGIK 166


RESULT 12
US-09-851-026-11
; Sequence 11, Application US/09851026
; Patent No. 6734168
;  GENERAL INFORMATION:
;       APPLICANT: Stern, David M.
;                  Clauss, Matthias
;                  Kao, Janet
;                  Kayton, Mark
;                  Libutti; Steven K
;       TITLE OF INVENTION: Endothelial Monocyte Activating
;                           Polypeptide II:  A Mediator Which Activates Host Response
;       NUMBER OF SEQUENCES: 42
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Cooper & Dunham, LLP
;            STREET: 1185 Avenue of the Americas
;            CITY: New York
;            STATE: New York
;            COUNTRY: USA
;            ZIP: 10036
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.30, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/851,026
;            FILING DATE: 07-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/360,821
;            FILING DATE: 08-OCT-96
;       ATTORNEY/AGENT INFORMATION:
;            NAME: White, John P.
;            REGISTRATION NUMBER: 28,678
;            REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-278-0400
;            TELEFAX: 212-391-0525
;   INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 166 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: Peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-851-026-11

  Query Match             53.1%;  Score 835;  DB 4;  Length 166;
  Best Local Similarity   94.0%;  Pred. No. 3.6e-69;
  Matches  156;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        147 SKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRM 206
              ||||| |||||||||||:||:||||||||||||||||| ||||||||||||||||||||||
Db          1 SKPIDASRLDLRIGCIVTAKKHPDADSLYVEEVDVGEAAPRTVVSGLVNHVPLEQMQNRM 60

Qy        207 VILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNP 266
              |:|||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||
Db         61 VVLLCNLKPAKMRGVLSQAMVMCASSPDKVEILAPPNGSVPGDRITFDAFPGEPDKELNP 120

Qy        267 KKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              ||||||||||||||| |||||||| ||||||||||||||:|||||
Db        121 KKKIWEQIQPDLHTNAECVATYKGAPFEVKGKGVCRAQTMANSGIK 166


RESULT 13
US-09-392-772-10
; Sequence 10, Application US/09392772
; Patent No. 6346403
; GENERAL INFORMATION:
;  APPLICANT: Famodu, Layo O.
;  APPLICANT: Antoni Rafalski
;  TITLE OF INVENTION: Methionine Metabolic Enzymes
;  FILE REFERENCE: BB-1241
;  CURRENT APPLICATION NUMBER: US/09/392,772
;  CURRENT FILING DATE: 1999-09-07
;  EARLIER APPLICATION NUMBER: 60/099,519
;  EARLIER FILING DATE: 1998-09-08
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
;   LENGTH: 659
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-392-772-10

  Query Match             32.5%;  Score 510.5;  DB 3;  Length 659;
  Best Local Similarity   38.6%;  Pred. No. 2.1e-38;
  Matches  123;  Conservative   49;  Mismatches   84;  Indels   63;  Gaps   11;

Qy         18 ADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKE----------------IEE 61
              :|:::    :|::|  |: :   :| |||    | || |                  :|
Db        380 SDEVL---RQLNLSPEENL---SLSEEKG---EIAKAKTPWNFVPAGHRIGKPAPLFKE 429

Qy         62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGD 119
              || |                    ||          :|:     :   |   |:||
Db        430 LKDE----------------DVALHREKYAGSQAERSSKAAADAEANKVANQLKG---- 468

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLY 175
                  |:   | ||| |:|| | : ||     | |::||:|:| |  | |||||||||
Db        469 -----TKLSDGGTKKEPKKQS--GGSKSKTAEADITVAKLDIRVGLIRKAEKHPDADSLY 521

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:|||| |||||||||  :|||:||:| | :||||||  |||: | |||: ||: :
Db        522 VEEIDVGEDAPRTVVSGLVKFIPLEEMQSRKVCVLCNLKPVAMRGIKSHAMVLAASNEDH 581

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
               |:|:: ||  :  |:|:||  : |||:  |: | | ||::  :||:| | || |: |||
Db        582 TKVELVEPPESAAVGERVTFAGYAGEPEASLSGKSKTWEKLAAELHSNGELVACYRDVPF 641

Qy        294 EVKGKGVCRAQTMSNSGIK 312
                  |||| :|::|  |:
Db        642 TTSA-GVCRVKTIANGEIR 659


RESULT 14
US-08-705-868-1
; Sequence 1, Application US/08705868
; Patent No. 5885798
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL MONOCYTE ACTIVATING CYTOKINE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/705,868
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0117 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  301 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:
;      CLONE:  Consensus
US-08-705-868-1

  Query Match             28.0%;  Score 440.5;  DB 2;  Length 301;
  Best Local Similarity   49.5%;  Pred. No. 1.9e-32;
  Matches   94;  Conservative   36;  Mismatches   51;  Indels    9;  Gaps    6;

Qy        123 AKEKIEKKGEKKEKKQQSIA-GSA-DSKPIDV--SRLDLRIGCIITARKHPDADSLYVEE 178
              | :|:          ||: :| | | :|:| :|  ||||:|:| |||  ||||||||||:
Db        105 ALKKLASAAYPDPSKQKPMAKGPAKNSEPEEVIPSRLDIRVGKIITVEKHPDADSLYVEK 164

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSP---EK 235
              :||||  |||||||||  || |::|:|:|::|||||| ||||| || |::|||      :
Db        165 IDVGEAEPRTVVSGLVQFVPKEELQDRLVVVLCNLKPQKMRGVESQGMLLCASIEGINRQ 224

Qy        236 IEILAPPNGSVPGDRITFDAF-PGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFE 294
              :| | || || ||: :    :  |:||:|| ||||::|::| |   ::||:| :|   |
Db        225 VEPLDPPAGSAPGEHVFVKGYEKGQPDEELKPKKKVFEKLQADFKISEECIAQWKQTNFM 284

Qy        295 VK-GKGVCRA 303
               | |   |::
Db        285 TKLGSISCKS 294


RESULT 15
US-09-123-615-1
; Sequence 1, Application US/09123615
; Patent No. 6090377
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL MONOCYTE ACTIVATING CYTOKINE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/123,615
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,868
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0117 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  301 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:
;      CLONE:  Consensus
US-09-123-615-1

  Query Match             28.0%;  Score 440.5;  DB 3;  Length 301;
  Best Local Similarity   49.5%;  Pred. No. 1.9e-32;
  Matches   94;  Conservative   36;  Mismatches   51;  Indels    9;  Gaps    6;

Qy        123 AKEKIEKKGEKKEKKQQSIA-GSA-DSKPIDV--SRLDLRIGCIITARKHPDADSLYVEE 178
              | :|:          ||: :| | | :|:| :|  ||||:|:| |||  ||||||||||:
Db        105 ALKKLASAAYPDPSKQKPMAKGPAKNSEPEEVIPSRLDIRVGKIITVEKHPDADSLYVEK 164

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSP---EK 235
              :||||  |||||||||  || |::|:|:|::|||||| ||||| || |::|||      :
Db        165 IDVGEAEPRTVVSGLVQFVPKEELQDRLVVVLCNLKPQKMRGVESQGMLLCASIEGINRQ 224

Qy        236 IEILAPPNGSVPGDRITFDAF-PGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFE 294
              :| | || || ||: :    :  |:||:|| ||||::|::| |   ::||:| :|   |
Db        225 VEPLDPPAGSAPGEHVFVKGYEKGQPDEELKPKKKVFEKLQADFKISEECIAQWKQTNFM 284

Qy        295 VK-GKGVCRA 303
               | |   |::
Db        285 TKLGSISCKS 294


Search completed: January 14, 2005, 15:59:58
Job time : 41 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 15:52:33 ; Search time 151 Seconds
(without alignments)
746.505 Million cell updates/sec

Title: US-09-733-306C-4
Perfect score: 1573
Sequence: 1 MANNDAVLKRLEQKGAEADQ.....PEVKGKGVCRAQTMSNSGIK 312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1568	99.7	312	9	US-09-851-026-36	Sequence 36, Appl
2	1343	85.4	310	9	US-09-851-026-35	Sequence 35, Appl
3	835	53.1	166	9	US-09-851-026-11	Sequence 11, Appl
4	831	52.8	166	10	US-09-930-169-3	Sequence 3, Appli
5	831	52.8	166	16	US-10-823-730-3	Sequence 3, Appli
6	709	45.1	147	10	US-09-930-169-1	Sequence 1, Appli
7	709	45.1	147	16	US-10-823-730-1	Sequence 1, Appli
8	528	33.6	353	15	US-10-424-599-180396	Sequence 180396,
9	517.5	32.9	377	17	US-10-425-115-350232	Sequence 350232,
10	517	32.9	804	16	US-10-437-963-183001	Sequence 183001,
11	516	32.8	108	10	US-09-930-169-2	Sequence 2, Appli
12	516	32.8	108	16	US-10-823-730-2	Sequence 2, Appli
13	511	32.5	792	16	US-10-437-963-183003	Sequence 183003,
14	510.5	32.5	810	15	US-10-425-114-37579	Sequence 37579, A
15	506.5	32.2	802	17	US-10-425-115-350231	Sequence 350231,
16	497.5	31.6	192	16	US-10-767-701-36700	Sequence 36700, A
17	451	28.7	536	9	US-09-813-718-2	Sequence 2, Appli
18	451	28.7	536	15	US-10-240-532-2	Sequence 2, Appli
19	451	28.7	536	16	US-10-240-527A-2	Sequence 2, Appli
20	450	28.6	528	17	US-10-370-715B-372	Sequence 372, App
21	440.5	28.0	301	10	US-09-919-039-132	Sequence 132, App
22	430.5	27.4	178	9	US-09-813-718-6	Sequence 6, Appli
23	430.5	27.4	178	15	US-10-240-532-6	Sequence 6, Appli
24	430.5	27.4	178	16	US-10-240-527A-6	Sequence 6, Appli
25	412	26.2	299	15	US-10-424-599-157861	Sequence 157861,
26	407.5	25.9	423	16	US-10-767-701-44089	Sequence 44089, A
27	402	25.6	409	16	US-10-437-963-109255	Sequence 109255,
28	389.5	24.8	351	17	US-10-425-115-197042	Sequence 197042,
29	378	24.0	267	15	US-10-425-114-53595	Sequence 53595, A
30	378	24.0	267	17	US-10-425-115-327897	Sequence 327897,
31	370.5	23.6	141	17	US-10-425-115-361131	Sequence 361131,
32	369	23.5	256	15	US-10-424-599-240858	Sequence 240858,
33	366	23.3	254	16	US-10-767-701-41506	Sequence 41506, A
34	337.5	21.5	185	16	US-10-437-963-163246	Sequence 163246,
35	315.5	20.1	217	17	US-10-425-115-283649	Sequence 283649,
36	293	18.6	58	9	US-09-813-718-22	Sequence 22, Appl
37	293	18.6	58	15	US-10-240-532-22	Sequence 22, Appl
38	293	18.6	58	16	US-10-240-527A-22	Sequence 22, Appl
39	291.5	18.5	364	17	US-10-425-115-304067	Sequence 304067,
40	253	16.1	158	17	US-10-425-115-197045	Sequence 197045,
41	233	14.8	307	15	US-10-424-599-157860	Sequence 157860,
42	225	14.3	690	15	US-10-282-122A-44568	Sequence 44568, A
43	197	12.5	109	15	US-10-424-599-202020	Sequence 202020,
44	193.5	12.3	113	17	US-10-425-115-287402	Sequence 287402,
45	188.5	12.0	669	15	US-10-282-122A-57809	Sequence 57809, A

ALIGNMENTS

```
RESULT 1
US-09-851-026-36
; Sequence 36, Application US/09851026
; Patent No. US20020160957A1
;   GENERAL INFORMATION:
;       APPLICANT: Stern, David M.
;                  Clauss, Matthias
;                  Kao, Janet
;                  Kayton, Mark
;                  Libutti, Steven K
;       TITLE OF INVENTION: Endothelial Monocyte Activating
;                           Polypeptide II:  A Mediator Which Activates Host Response
;       NUMBER OF SEQUENCES: 42
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Cooper & Dunham, LLP
;            STREET: 1185 Avenue of the Americas
;            CITY: New York
;            STATE: New York
;            COUNTRY: USA
;            ZIP: 10036
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.30, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/851,026
;            FILING DATE: 07-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/360,821
;            FILING DATE: 08-OCT-96
;       ATTORNEY/AGENT INFORMATION:
;            NAME: White, John P.
;            REGISTRATION NUMBER: 28,678
;            REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
```

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 212-278-0400
;              TELEFAX: 212-391-0525
;    INFORMATION FOR SEQ ID NO: 36:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 312 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-851-026-36

  Query Match             99.7%;  Score 1568;  DB 9;  Length 312;
  Best Local Similarity   99.7%;  Pred. No. 4.4e-120;
  Matches  311;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIPFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        301 CRAQTMSNSGIK 312


RESULT 2
US-09-851-026-35
; Sequence 35, Application US/09851026
; Patent No. US20020160957A1
;   GENERAL INFORMATION:
;        APPLICANT: Stern, David M.
;                   Clauss, Matthias
;                   Kao, Janet
;                   Kayton, Mark
;                   Libutti, Steven K
;        TITLE OF INVENTION: Endothelial Monocyte Activating
;                            Polypeptide II:  A Mediator Which Activates Host Response
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Cooper & Dunham, LLP
;             STREET: 1185 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.30, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/851,026
;             FILING DATE: 07-May-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/360,821
;             FILING DATE: 08-OCT-96
;        ATTORNEY/AGENT INFORMATION:
;             NAME: White, John P.
;             REGISTRATION NUMBER: 28,678
;             REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-278-0400
;             TELEFAX: 212-391-0525
;    INFORMATION FOR SEQ ID NO: 35:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 310 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-851-026-35

  Query Match             85.4%;  Score 1343;  DB 9;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 1.2e-101;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:|||||||||||:||||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     ||| |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| ||||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              |||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 3
US-09-851-026-11
; Sequence 11, Application US/09851026
; Patent No. US20020160957A1
;   GENERAL INFORMATION:
;        APPLICANT: Stern, David M.
;                   Clauss, Matthias
;                   Kao, Janet
;                   Kayton, Mark
;                   Libutti, Steven K
;        TITLE OF INVENTION: Endothelial Monocyte Activating
;                            Polypeptide II:  A Mediator Which Activates Host Response
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Cooper & Dunham, LLP
;             STREET: 1185 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.30, Version #1.30
;        CURRENT APPLICATION DATA:
```

```
;              APPLICATION NUMBER: US/09/851,026
;              FILING DATE: 07-May-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US/08/360,821
;              FILING DATE: 08-OCT-96
;         ATTORNEY/AGENT INFORMATION:
;              NAME: White, John P.
;              REGISTRATION NUMBER: 28,678
;              REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 212-278-0400
;              TELEFAX: 212-391-0525
;    INFORMATION FOR SEQ ID NO: 11:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 166 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-851-026-11

  Query Match             53.1%;  Score 835;  DB 9;  Length 166;
  Best Local Similarity   94.0%;  Pred. No. 2.1e-60;
  Matches  156;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        147 SKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRM 206
              ||||| |||||||||||:||:||||||||||||||| ||||||||||||||||||||||
Db          1 SKPIDASRLDLRIGCIVTAKKHPDADSLYVEEVDVGEAAPRTVVSGLVNHVPLEQMQNRM 60

Qy        207 VILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNP 266
              |:|||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||
Db         61 VVLLCNLKPAKMRGVLSQAMVMCASSPDKVEILAPPNGSVPGDRITFDAFPGEPDKELNP 120

Qy        267 KKKIWEQIQPDLHTNDECVATYKGVPPEVKGKGVCRAQTMSNSGIK 312
              ||||||||||||||| |||||||| |||||||||||||||:||||||
Db        121 KKKIWEQIQPDLHTNAECVATYKGAPPEVKGKGVCRAQTMANSGIK 166


RESULT 4
US-09-930-169-3
; Sequence 3, Application US/09930169
; Publication No. US20030004309A1
; GENERAL INFORMATION:
;  APPLICANT: KIM, SUNGHOON
;  APPLICANT:  KO, YOUNG-GYU
;  TITLE OF INVENTION: IMMUNOLOGICAL ENHANCEMENT AGENT COMPRISING N-TERMINAL PEPTIDE OF
;  TITLE OF INVENTION:  P43 AS AN EFFECTIVE COMPONENT
;  FILE REFERENCE: 058333/0106
;  CURRENT APPLICATION NUMBER: US/09/930,169
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: KR 2001-31310
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian
;   OTHER INFORMATION: protein sequence
US-09-930-169-3

  Query Match             52.8%;  Score 831;  DB 10;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-60;
  Matches  166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         91 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 60

Qy        151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 120

Qy        211 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 256
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 166


RESULT 5
US-10-823-730-3
; Sequence 3, Application US/10823730
; Publication No. US20040185060A1
; GENERAL INFORMATION:
;  APPLICANT: IMAGENE CO., LTD.
;  TITLE OF INVENTION: Immunological enhancement agent comprising N-terminal peptide of
;  TITLE OF INVENTION:  p43 as an effective component
;  FILE REFERENCE: NPF1918
;  CURRENT APPLICATION NUMBER: US/10/823,730
;  CURRENT FILING DATE:  2004-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: KopatentIn 1.71
; SEQ ID NO 3
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: mammalian
US-10-823-730-3

  Query Match             52.8%;  Score 831;  DB 16;  Length 166;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-60;
  Matches  166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         91 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPI 60

Qy        151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 120

Qy        211 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 256
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CNLKPAKMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAF 166


RESULT 6
US-09-930-169-1
; Sequence 1, Application US/09930169
; Publication No. US20030004309A1
; GENERAL INFORMATION:
;  APPLICANT: KIM, SUNGHOON
;  APPLICANT:  KO, YOUNG-GYU
;  TITLE OF INVENTION: IMMUNOLOGICAL ENHANCEMENT AGENT COMPRISING N-TERMINAL PEPTIDE OF
;  TITLE OF INVENTION:  P43 AS AN EFFECTIVE COMPONENT
;  FILE REFERENCE: 058333/0106
;  CURRENT APPLICATION NUMBER: US/09/930,169
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: KR 2001-31310
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 147
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian
;   OTHER INFORMATION: protein sequence
US-09-930-169-1

  Query Match             45.1%;  Score 709;  DB 10;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-50;
  Matches  147;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147
              ||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147
```

RESULT 7
US-10-823-730-1
; Sequence 1, Application US/10823730
; Publication No. US20040185060A1
; GENERAL INFORMATION:
; APPLICANT: IMAGENE CO., LTD.
; TITLE OF INVENTION: Immunological enhancement agent comprising N-terminal peptide of
; TITLE OF INVENTION: p43 as an effective component
; FILE REFERENCE: NPF1918
; CURRENT APPLICATION NUMBER: US/10/823,730
; CURRENT FILING DATE: 2004-04-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 1
; LENGTH: 147
; TYPE: PRT
; ORGANISM: mammalian
US-10-823-730-1

Query Match 45.1%; Score 709; DB 16; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.8e-50;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147
              ||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADS 147
```

RESULT 8
US-10-424-599-180396
; Sequence 180396, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 180396
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133912C.1.pep
US-10-424-599-180396

Query Match 33.6%; Score 528; DB 15; Length 353;
Best Local Similarity 43.6%; Pred. No. 8.6e-35;
Matches 113; Conservative 41; Mismatches 79; Indels 26; Gaps 4;

```
Qy         56 KKEIEELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKG 115
              |    ||| |    |:.   |: |      :      ::|| :    | || |  :: |
Db        119 KPLFRELKDE----EVEFYRKKFAGSQADRIVRAEAEAQNVAEQLKKTKVSDGNGKKKPG 174

Qy        116 GTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLY 175
              : :| | |                    :|    | ::|||:|:| || |:||||:||
Db        175 KSSNEAKNK------------------AAAEPDITITRLDIRVGLIIKAQKHPDADALY 215

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:||||   |||||||| :||::|||| | :|||||  |||: |||||: ||  :
Db        216 VEEIDVGEEQTRTVVSGLVKFIPLDEMQNRKVCVLCNLKPVTMRGIKSQAMVLAASDGDH 275

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
               |:|:: ||: : ||:|||| : | ||: ||||||:|| :| |||||:| || || ||
Db        276 TKVELVEPPSSAQPGERITFPGYEGNPDELLNPKKKVWETLQVDLHTNEELVACYKNVPL 335

Qy        294 EVKGKGVCRAQTMSNSGIK 312
                |||:  ::|   |:
Db        336 TTSA-GVCKVSSISCGSIR 353
```

RESULT 9
US-10-425-115-350232
; Sequence 350232, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 350232
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_82583C.1.pep
US-10-425-115-350232

Query Match 32.9%; Score 517.5; DB 17; Length 377;
Best Local Similarity 38.2%; Pred. No. 6.8e-34;
Matches 122; Conservative 50; Mismatches 84; Indels 63; Gaps 11;

```
Qy         18 ADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKE----------------IEE 61
              :|:::   :|::|  |: :  :| |||    | || |                    :|
Db         98 SDEVL----RQLNLSPEENL---SLSEEKG---EIAKAKSPWDFVPAGHRIGKPAPLFKE 147

Qy         62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGD 119
              || |                ||          :|:     :    |   |:||
Db        148 LKDE----------------DVALHREKYAGSQAERSSKAAADAEANKVANQLKG---- 186

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPID----VSRLDLRIGCIITARKHPDADSLY 175
                 |:  :| ||| |:||  | : ||  |    |::||:|:| |  | |||||||||
Db        187 -----TKLSDRGTKKEPKKQS--GGSKSKTADADITVAKLDIRVGLIRKAEKHPDADSLY 239

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:|||| |||||||||  :|||:|||| | :|||||  |||: | |||: ||: :
Db        240 VEEIDVGEDTPRTVVSGLVKFIPLEEMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDH 299

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
```

```
Db        300 TKVELVEPPESAAVGERVTFAGYSGEPEASLSGKSKTWEKLAAELHSNGELVACYKDVPF 359

Qy        294 EVKGKGVCRAQTMSNSGIK 312
Db        360 TTSA-GICKVKTIANGEIR 377


RESULT 10
US-10-437-963-183001
; Sequence 183001, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183001
;   LENGTH: 804
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_80132C.1.pep
US-10-437-963-183001

  Query Match             32.9%;  Score 517;  DB 16;  Length 804;
  Best Local Similarity   38.1%;  Pred. No. 2.1e-33;
  Matches  117;  Conservative   53;  Mismatches   79;  Indels   58;  Gaps    9;

Qy         35 KAILQATLREEKKL-----RVENAKLKKE----------------IEELKQELIQAEIQN 73
Db        527 EVLLQLNMTPEESLSFCDDKGEIAKAKRPWDFVSAGHKIGKPSPLFKELKDEEVES---- 582

Qy         74 GVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGDEKKAKEKIEKKG 131
Db        583 -------------FRNKFAGSQAERSSKAQADAEAKKVADKLKG---------TKLSDGG 620

Qy        132 EKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPR 187
Db        621 QKKEQKKQS--GGSKSKNAEVDVTVAKLDIRVGLIRKAQKHPDADSLYVEEIDVGEEAPR 678

Qy        188 TVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAPPNGS 245
Db        679 TVVSGLVKFIPLEEMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDHTKVELVEPPESA 738

Qy        246 VPGDRITFDAPPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQT 305
Db        739 AVGERVTFAGYSGEPEASLNAKSKTWEKLSADLHSNGELVACYKDVPFTTSA-GVCKVKS 797

Qy        306 MSNSGIK 312
Db        798 IASGEIR 804


RESULT 11
US-09-930-169-2
; Sequence 2, Application US/09930169
; Publication No. US20030004309A1
; GENERAL INFORMATION:
;  APPLICANT: KIM, SUNGHOON
;  APPLICANT:  KO, YOUNG-GYU
;  TITLE OF INVENTION: IMMUNOLOGICAL ENHANCEMENT AGENT COMPRISING N-TERMINAL PEPTIDE OF
;  TITLE OF INVENTION:  P43 AS AN EFFECTIVE COMPONENT
;  FILE REFERENCE: 058333/0106
;  CURRENT APPLICATION NUMBER: US/09/930,169
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: KR 2001-31310
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 108
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian
;   OTHER INFORMATION: protein sequence
US-09-930-169-2

  Query Match             32.8%;  Score 516;  DB 10;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-34;
  Matches  108;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108


RESULT 12
US-10-823-730-2
; Sequence 2, Application US/10823730
; Publication No. US20040185060A1
; GENERAL INFORMATION:
;  APPLICANT: IMAGENE CO., LTD.
;  TITLE OF INVENTION: Immunological enhancement agent comprising N-terminal peptide of
;  TITLE OF INVENTION:  p43 as an effective component
;  FILE REFERENCE: NPF1918
;  CURRENT APPLICATION NUMBER: US/10/823,730
;  CURRENT FILING DATE:  2004-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: KopatentIn 1.71
; SEQ ID NO 2
;   LENGTH: 108
;   TYPE: PRT
;   ORGANISM: mammalian
US-10-823-730-2

  Query Match             32.8%;  Score 516;  DB 16;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-34;
  Matches  108;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSG 108


RESULT 13
US-10-437-963-183003
; Sequence 183003, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
```

```
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183003
;   LENGTH: 792
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_80134C.1.pep
US-10-437-963-183003

  Query Match            32.5%;  Score 511;  DB 16;  Length 792;
  Best Local Similarity  35.9%;  Pred. No. 6.4e-33;
  Matches  127;  Conservative   64;  Mismatches   99;  Indels   64;  Gaps   13;

Qy          9 KRLEQ--KGAEADQIIEYLKQQVSLLKE-----KAILQATLREEKKL-----RVENAKLK 56
              | :||  :  |  :: : ||  : :  :      : :||  :  |: |     : | || |
Db        453 KWVEQYLEAMEKVKLKQGLKSAMGISSDGNAYLQVLLQLNMTPEESLSFCDDKGEIAKAK 512

Qy         57 KE---------------IEELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQST 100
              :                :||| |    |:::   : |   |:    :|    :  ::
Db        513 RPWDFVSAGHKIGKPSPLFKELKDE----EVESFRNKFA---GSQAERSSKAQADA-EAK 564

Qy        101 AVTTVSSGTKEQIKGGTGDEKKAKEKIE---------------KKGEKKEKKQQSIAGS 144
              |     |||      | | ::   ::                    |:|||:|:||  |
Db        565 KVADKLKGTK---LSGENDWEQKNRSLQVGNEISRVGNLDYWETDGGQKKEQKKQS--GG 619

Qy        145 ADSK----PIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLE 200
              : ||    : |::||:|:| |  |:||||||||||||:|||| ||||||||||  :|||
Db        620 SKSKNAEVDVTVAKLDIRVGLIRKAQKHPDADSLYVEEIDVGEEAPRTVVSGLVKFIPLE 679

Qy        201 QMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAPPNGSVPGDRITFDAFPG 258
              :|||| | :|||||  |||: | |||: ||: :  |:|:: ||   :  |:|:||  : |
Db        680 EMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDHTKVELVEPPESAAVGERVTFAGYSG 739

Qy        259 EPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGIK 312
              ||:  || | | ||::  |||:| | || || |||     |||: :::::  |:
Db        740 EPEASLNAKSKTWEKLSADLHSNGELVACYKDVPFTTSA-GVCKVKSIASGEIR 792


RESULT 14
US-10-425-114-37579
; Sequence 37579, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37579
;   LENGTH: 810
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3689-210-B11_FLI.pep
US-10-425-114-37579

  Query Match            32.5%;  Score 510.5;  DB 15;  Length 810;
  Best Local Similarity  38.6%;  Pred. No. 7.2e-33;
  Matches  123;  Conservative   49;  Mismatches   84;  Indels   63;  Gaps   11;

Qy         18 ADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKE---------------IEE 61
              :|:::    :|::|  |: :  :| |||    | || |                 :|
Db        531 SDEVL----RQLNLSPEENL---SLSEEKG---EIAKAKTPWNFVPAGHRIGKPAPLFKE 580

Qy         62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGD 119
              || |                 ||         :|:     :   |   |:||
Db        581 LKDE----------------DVALHREKYAGSQAERSSKAAADAEANKVANQLKG---- 619

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLY 175
                    |:   | ||| |:||  | : ||      | |::||:|:|  |  | |||||||||
Db        620 -----TKLSDGGTKKEPKKQS--GGSKSKTAEADITVAKLDIRVGLIRKAEKHPDADSLY 672

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:|||| |||||||||  :|||:||:| | :||||||  |||: | |||: ||: :
Db        673 VEEIDVGEDAPRTVVSGLVKFIPLEEMQSRKVCVLCNLKPVAMRGIKSHAMVLAASNEDH 732

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
               |:|:: ||  :  |:|:||  : |||:  |: | | ||::  :||:| | || |: |||
Db        733 TKVELVEPPESAAVGERVTFAGYAGEPEASLSGKSKTWEKLAAELHSNGELVACYRDVPF 792

Qy        294 EVKGKGVCRAQTMSNSGIK 312
                   |||| :|::|  |:
Db        793 TTSA-GVCRVKTIANGEIR 810


RESULT 15
US-10-425-115-350231
; Sequence 350231, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 350231
;   LENGTH: 802
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_82582C.1.pep
US-10-425-115-350231

  Query Match            32.2%;  Score 506.5;  DB 17;  Length 802;
  Best Local Similarity  38.2%;  Pred. No. 1.5e-32;
  Matches  122;  Conservative   49;  Mismatches   85;  Indels   63;  Gaps   11;

Qy         18 ADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKE---------------IEE 61
              :|:::    :|::|  |: :  :| |||    | || |                 :|
Db        523 SDEVL----RQLNLSPEENL---SLSEEKG---EIAKAKTPWNFVPAGHRIGKPAPLFKE 572

Qy         62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGD 119
              || |                 ||         :|:     :   |   |:||
Db        573 LKDE----------------DVALHREKYAGSQAERSSKAAADAEANKVANQLKG---- 611

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLY 175
                    |:   | ||| |:||  | : ||      | |::||:|:|  |  | |||||||||
Db        612 -----TKLSDGGTKKEPKKQS--GGSKSKTAEADITVAKLDIRVGLIRKAEKHPDADSLY 664

Qy        176 VEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE- 234
              |||:|||| |||||||||  :|||:||:| | :||||||  |||: | |||: ||: :
Db        665 VEEIDVGEDTPRTVVSGLVKFIPLEEMQSRKVCVLCNLKPVAMRGIKSHAMVLAASNEDH 724

Qy        235 -KIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
               |:|:: ||  :  |:|:||  : |||:  |: | | ||::  :||:| | || |: |||
Db        725 TKVELVEPPESAAVGERVTFAGYAGEPEASLSGKSKTWEKLAAELHSNGELVACYRDVPF 784

Qy        294 EVKGKGVCRAQTMSNSGIK 312
```

```
                    |||| :|::|  |:
Db        785 TTSA-GVCRVKTIANGEIR 802
```

Search completed: January 14, 2005, 16:02:35
Job time : 153 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 15:49:22 ; Search time 40 Seconds
(without alignments)
750.491 Million cell updates/sec

Title: US-09-733-306C-4
Perfect score: 1573
Sequence: 1 MANNDAVLKRLEQKGAEADQ.....PEVKGKGVCRAQTMSNSGIK 312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1546	98.3	310	2	B55053	endothelial monocy
2	1343	85.4	310	2	A55053	endothelial monocy
3	564	35.9	917	2	T22898	hypothetical prote
4	499	31.7	797	2	T05247	methionine-tRNA li
5	429	27.3	376	2	S64113	ARC1 protein - yea
6	400.5	25.5	440	2	C84832	probable methionyl
7	343	21.8	273	2	T47822	hypothetical prote
8	223.5	14.2	722	2	B75074	methionyl-tRNA syn
9	216	13.7	723	2	D71091	methionine-tRNA li
10	205	13.0	616	1	SYTWMT	methionine-tRNA li
11	198.5	12.6	110	2	C85968	probable tRNA synt
12	198.5	12.6	110	2	D91123	probable tRNA synt
13	194.5	12.4	110	2	G65095	hypothetical 12.3
14	191	12.1	658	2	D69431	methionine-tRNA li
15	185.5	11.8	629	2	E72297	methionine-tRNA li
16	183.5	11.7	650	2	A64572	methionine-tRNA li
17	182.5	11.6	651	2	F64457	methionine-tRNA li
18	178.5	11.3	811	2	E71281	methionine-tRNA li
19	177	11.3	660	2	E83656	methionyl-tRNA syn
20	170	10.8	656	2	H71867	methionine-tRNA li
21	168.5	10.7	114	2	F82090	methionyl-tRNA syn
22	168.5	10.7	681	2	C75395	methionyl-tRNA syn
23	168	10.7	734	2	B70173	methionine-tRNA li
24	159.5	10.1	664	2	AI1459	methionyl-tRNA syn
25	158	10.0	664	2	S66067	methionine-tRNA li
26	156.5	9.9	665	2	D95091	methionyl-tRNA syn
27	156.5	9.9	679	2	H97958	methionine-tRNA li
28	155.5	9.9	664	2	AB1097	methionyl-tRNA syn
29	151	9.6	644	2	A97268	methionyl-tRNA syn
30	149	9.5	651	2	G69177	methionine-tRNA li
31	147	9.3	111	2	A83243	CsaA protein PA322
32	144	9.2	649	2	S16682	methionine-tRNA li
33	142.5	9.1	111	2	F70338	methionine-tRNA li
34	141.5	9.0	117	2	H72655	protein secretion
35	141	9.0	682	2	I64113	methionine-tRNA li
36	139.5	8.9	677	2	AH0776	methionine-tRNA li
37	135	8.6	662	2	H86723	methionine-tRNA li
38	134.5	8.6	628	2	F81356	methionine-tRNA li
39	134.5	8.6	677	2	C83210	methionyl-tRNA syn
40	133	8.5	677	2	C85839	methionine tRNA sy
41	133	8.5	677	2	H90993	methionine tRNA sy
42	133	8.5	731	2	D82249	methionyl-tRNA syn
43	132.5	8.4	677	1	SYECMT	methionine-tRNA li
44	130	8.3	276	2	T23451	hypothetical prote
45	129.5	8.2	689	2	C84192	methionine-tRNA sy

ALIGNMENTS

RESULT 1
B55053
endothelial monocyte-activating protein II precursor - human
C;Species: Homo sapiens (man)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Dec-2002
C;Accession: B55053
R;Kao, J.; Houck, K.; Fan, Y.; Haehnel, I.; Libutti, S.K.; Kayton, M.L.; Grikscheit, T.
, J.; Stern, D.M.
J. Biol. Chem. 269, 25106-25119, 1994
A;Title: Characterization of a novel tumor-derived cytokine. Endothelial-monocyte activ
A;Reference number: A55053; MUID:95014290; PMID:7929199
A;Accession: B55053
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-310 <KAO>
A;Cross-references: GB:U10117
C;Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)

Query Match 98.3%; Score 1546; DB 2; Length 310;
Best Local Similarity 99.0%; Pred. No. 1.9e-87;
Matches 309; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              |||||||||||||||||| |||||||||||||||||||||||||||||||  |||
Db         61 ELKQELIQAEIQNGVKQIRFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKG--GDE 118

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 178

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 238

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 298

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        299 CRAQTMSNSGIK 310
```

RESULT 2
A55053
endothelial monocyte-activating protein II precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

```
C;Accession: A55053; A44032
R;Kao, J.; Houck, K.; Fan, Y.; Haehnel, I.; Libutti, S.K.; Kayton, M.L.; Grikscheit, T.;
, J.; Stern, D.M.
J. Biol. Chem. 269, 25106-25119, 1994
A;Title: Characterization of a novel tumor-derived cytokine. Endothelial-monocyte activa
A;Reference number: A55053; MUID:95014290; PMID:7929199
A;Accession: A55053
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-310 <KAO>
A;Cross-references: UNIPROT:P31230; GB:U10118; NID:g498911; PIDN:AAA62203.1; PID:g498912
R;Kao, J.; Ryan, J.; Brett, G.; Chen, J.; Shen, H.; Fan, Y.G.; Godman, G.; Familletti, P
J. Biol. Chem. 267, 20239-20247, 1992
A;Title: Endothelial monocyte-activating polypeptide II. A novel tumor-derived polypepti
A;Reference number: A44032; MUID:93015897; PMID:1400342
A;Accession: A44032
A;Molecule type: protein
A;Residues: 145-158,'X',160-164 <KA2>
A;Experimental source: methylcholanthrene A fibrosarcoma cells
A;Note: sequence extracted from NCBI backbone (NCBIP:115676)
C;Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)

  Query Match             85.4%;  Score 1343;  DB 2;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 4.7e-75;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:||||||||||:||||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| ||:|:     ||| |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:|||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| ||||||||||||||||||||||:|||||||||||||||||||||||||||:|||
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              |||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 3
T22898
hypothetical protein F58B3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22898
R;Harris, B.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19633
A;Accession: T22898
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-917 <WIL>
A;Cross-references: UNIPROT:Q20970; EMBL:Z73427; PIDN:CAA97803.1; GSPDB:GN00022; CESP:F5
A;Experimental source: clone F58B3
C;Genetics:
A;Gene: CESP:F58B3.5
A;Map position: 4
A;Introns: 27/2; 135/2; 620/1; 655/1; 874/3
C;Superfamily: methionyl-tRNA synthetase, dimer-forming

  Query Match             35.9%;  Score 564;  DB 2;  Length 917;
  Best Local Similarity   42.3%;  Pred. No. 6e-27;
  Matches  135;  Conservative   52;  Mismatches  100;  Indels   32;  Gaps    8;

Qy          1 MANNDAVLKRLEQKGAEADQII-EYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEI 59
              ||:  | :: |||       |:| : ||:     : ||:   |  : ::|  || :|  ::
Db        624 MASTAAFVE-LEQGAKVISQLIAQNLKK---FDQAKALF--TRNQLQRLDGENKQLTIDV 677

Qy         60 EELKQELIQAEIQNGVKQIAFP--SGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGT 117
              : |: :||: |   |:||:  |  | ||              |  :| :||   :
Db        678 KTLQHQLIELETAAGIKQVPKPVVSCTP-------------TPTSTPASGIITE----- 718

Qy        118 GDEKKAKEKIEKKGEKKEKKQQSIAGSADSKP----IDVSRLDLRIGCIITARKHPDADS 173
                :|:|          || |:|:   | | : |    ||| |||:|:| ||   ||||||:
Db        719 APKKEAPSTPAPSEPKKAKEQKKGKGGAAAAPVDDTIDVGRLDMRVGRIIKCEKHPDADA 778

Qy        174 LYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSP 233
              ||||::|||| |||||||||| ||||:|||||:|::|||||||||||| |:|||||||||
Db        779 LYVEQIDVGESAPRTVVSGLVRHVPLDQMQNRLVVVLCNLKPAKMRGVESRAMVMCASSP 838

Qy        234 EKIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
              :|:||:  |  | ||  :    :   ||::|||||||| :  ||  : |  | :|| |
Db        839 DKVEIMEVPADSKPGTPVVCPPYTHRPDEQLNPKKKIWETVAEDLKVSAEGFAEWKGQPL 898

Qy        294 EVKGKGVCRAQTMSNSGIK 312
              :  :    | |:    :|
Db        899 LIGSESKMTAPTLRGVHVK 917


RESULT 4
T05247
methionine-tRNA ligase homolog F18A5.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05247
R;Bevan, M.; Weber, N.; Grueninger, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15405
A;Accession: T05247
A;Molecule type: DNA
A;Residues: 1-797 <BEV>
A;Cross-references: UNIPROT:Q9SVN5; EMBL:AL035528
A;Experimental source: cultivar Columbia; BAC clone F18A5
C;Genetics:
A;Map position: 4
A;Introns: 40/1; 88/2; 139/3; 191/3; 268/3; 316/3; 347/1; 382/3; 423/1; 460/3; 482/3; 5
A;Note: F18A5.170
C;Superfamily: methionyl-tRNA synthetase, dimer-forming

  Query Match             31.7%;  Score 499;  DB 2;  Length 797;
  Best Local Similarity   52.9%;  Pred. No. 4.8e-23;
  Matches  101;  Conservative   31;  Mismatches   51;  Indels    8;  Gaps    3;

Qy        124 KEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGE 183
              |:|   ||  | | |     |  : | ::|||:|:|  |: | ||| ||:|||||:|||
Db        613 KQKASSKGGGKPKPQ-----PAADREITMARLDIRVGKIVKAEKHPKADALYVEEIDVGG 667

Qy        184 IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAP 241
              |||||||| ::|||:|||||| :|||||||||| ::|||||: ||| :  |:|:: |
Db        668 GEIRTVVSGLVKYIPLEEMQNRMVCVLCNLKPAKMRDIVSQAMVLAASSSDGSKVELVEP 727

Qy        242 PNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVC 301
              |  :  |:|:||  | ||||  ||||||:|| :  ||:|  :  || || |||     |||
Db        728 PKTANIGERVTFPGFEGEPDDVLNPKKKVWETLLVDLNTKENLVACYKDVPFTTSA-GVC 786

Qy        302 RAQTMSNSGIK 312
              :  ::||  |:
Db        787 KVSSISNGTIR 797


RESULT 5
S64113
ARC1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: G4P1 protein; protein G3085; protein YGL105w
```

C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
C;Accession: S64113; S69425; S72274
R;Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64112
A;Accession: S64113
A;Molecule type: DNA
A;Residues: 1-376 <CAS>
A;Cross-references: UNIPROT:P46672; EMBL:Z72627; NID:g1322647; PIDN:CAA96812.1; PID:g132
A;Experimental source: strain S288C
R;Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the EMBL Data Library, April 1996
A;Reference number: S69417
A;Accession: S69425
A;Molecule type: DNA
A;Residues: 1-376 <CAW>
A;Cross-references: EMBL:X97644; NID:g1310710; PIDN:CAA66247.1; PID:g1310719
R;Simos, G.; Segref, A.; Fasiolo, F.; Hellmuth, K.; Shevchenko, A.; Mann, M.; Hurt, E.C.
EMBO J. 15, 5437-5448, 1996
A;Title: The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methi
A;Reference number: S72274; MUID:97050848; PMID:8895587
A;Accession: S72274
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-376 <SIM>
A;Cross-references: EMBL:X95481; NID:g1620459; PIDN:CAA64750.1; PID:g1620460
C;Genetics:
A;Gene: SGD:ARC1; G4P1
A;Cross-references: MIPS:YGL105w; SGD:S0003073
A;Map position: 7L
C;Complex: homodimer
C;Function:
A;Description: functions as a cofactor for methionyl- and glutamyl-tRNA synthetases
C;Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)
C;Keywords: cytosol; homodimer

Query Match 27.3%; Score 429; DB 2; Length 376;
Best Local Similarity 42.7%; Pred. No. 3.7e-19;
Matches 105; Conservative 32; Mismatches 81; Indels 28; Gaps 9;

```
Qy          56 KKEIE---ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQ 112
               | ||    :|  |:|:       |:   |:|    | :   |:| :           |:
Db         116 KLEINHDLDLPHEVIE-------KKKKAPAGGAADAAAKADEDVSK--------KAKKQD 160

Qy         113 IKGGTGDE---KKAKEKIE-KKGEKK---EKKQQSIAGSADSKPIDVSRLDLRIGCIITA 165
                 |  ||    || :|: : ||  ||     |:||      |  ||    | :| |:| |  |
Db         161 HPRGKPDEETLKKLREEAKAKKAAKKAANAKQQQEQQNKAPEKP-KPSAIDFRVGFIQKA 219

Qy         166 RKHPDADSLYVEEVDVG-EIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQ 224
                |||||||||||  :||| |  |||| |||| | ||: || | |:::|||||  |||: |
Db         220 IKHPDADSLYVSTIDVGDEEGPRTVCSGLVKHFPLDAMQERYVVVVCNLKPVNMRGIKST 279

Qy         225 AMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGE-PDKELNPKKKIWEQIQPDLHTNDE 283
               |||:| |: :|:| : ||  |  ||:: |: |  | | |:|||||||| :||   |||
Db         280 AMVLCGSNDDKVEFVEPPKDSKAGDKVFFEGFGDEAPMKQLNPKKKIWEHLQPHFTTNDG 339

Qy         284 CVATYK 289
                  :|
Db         340 LEVIFK 345
```

RESULT 6
C84832
probable methionyl-tRNA synthetase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84832
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84832
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-440 <STO>
A;Cross-references: UNIPROT:Q7XJM9; GB:AE002093; NID:g4895232; PIDN:AAD32818.1; GSPDB:GN
C;Genetics:
A;Gene: At2g40660
A;Map position: 2
C;Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)

Query Match 25.5%; Score 400.5; DB 2; Length 440;
Best Local Similarity 35.6%; Pred. No. 2.5e-17;
Matches 105; Conservative 45; Mismatches 86; Indels 59; Gaps 9;

```
Qy          22 IEYLKQQVSLLKEKAILQATLRE------EKKLRVE-NAKLKKEIEELKQELIQAEIQNG 74
               : |:: :  |    | :   | |       :  ::|| |:  ||  |             |
Db         180 VNYIQNKEELSTLFAPIPVKLPEFSFEVPKPAIKVETNSNSKKAAE------------G 226

Qy          75 VKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKK 134
               || :  |   |                    |||:     | |:  |   ::|  ||
Db         227 VKPVDKPDVQP-------------------QLGTKK-----TEPEEPKKNAAKEKDAKK 261

Qy         135 EKKQQSIAGSADSKP------IDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRT 188
               |||:      |: :|      : || |:::| |  | ||| |||| |||:||||   |
Db         262 EKKK-----PAEPEPAKKEAELSVSLLNIQVGLIRKAWKHPSADSLLVEEIDVGEDKVRQ 316

Qy         189 VVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEK--IEILAPPNGSV 246
               |||||      | : ||:| |: |:|| |:| |:|| :|:|||| :   :| | || |:
Db         317 VVSGLAKFCSPEDLTNRLVALITNVKPGKLRDVMSQGLVLCASSEDHSVVEPLLPPAGAK 376

Qy         247 PGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVC 301
               ||:|::|    |:|:  ||||||  |:| | |:|::  ||||||: | :   | |
Db         377 PGERVSFSGIEGKPEDVLNPKKKQLEKITPGLYTDENGVATYKGIQF-MTSAGPC 430
```

RESULT 7
T47822
hypothetical protein F24G16.250 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 17-Feb-2003
C;Accession: T47822
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24477
A;Accession: T47822
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-273 <DAN>
A;Cross-references: EMBL:AL138647
A;Experimental source: cultivar Columbia; BAC clone F24G16
C;Genetics:
A;Map position: 3
A;Introns: 157/3; 226/3
A;Note: F24G16.250
C;Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)

Query Match 21.8%; Score 343; DB 2; Length 273;
Best Local Similarity 37.8%; Pred. No. 4.6e-14;
Matches 84; Conservative 40; Mismatches 58; Indels 40; Gaps 7;

```
Qy         101 AVTTVSSGTKEQIKGGTGDEKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIG 160
               | ||||:                  :: ||| |  :|:::::  :|       : ||:::|
Db          80 AGTTVSA----------------DESEKKSE-SQKEEENVKETA-------NLLDIKVG 114

Qy         161 CIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRG 220
                |: | :| :|||||||||||:||  || : |||| :|||: :|   |::| ||||  |||
Db         115 RIVKAWQHEEADSLYVEEVDIGEAEPRIICSGLVKYVPLDLLQGASVVVLANLKPRNMRG 174

Qy         221 VLSQAMVMCAS--SPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNP---------KKK 269
               | |  |:: ||  : | :|:| || |||||||: |     | | |  |         |||
Db         175 VKSCGMLLAASDAAHENVELLVPPEGSVPGDRVWFG---NEEDLEQLPEPAPPNKVQKKK 231
```

```
Qy        270 IWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQTMSNSGI 311
              :|| :|| | |:   |:  |          |:  :::: |: |
Db        232 MWELVQPLLKTDASGVSMLKEHLMRT-SSGLVTSKSLRNANI 272
```

RESULT 8
B75074
methionyl-tRNA synthetase (mets) PAB2364 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B75074
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: B75074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-722 <KAW>
A;Cross-references: UNIPROT:Q9V011; GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB4989
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: metS; PAB2364
C;Superfamily: methionyl-tRNA synthetase, dimer-forming

  Query Match 14.2%; Score 223.5; DB 2; Length 722;
  Best Local Similarity 29.7%; Pred. No. 2.8e-06;
  Matches 77; Conservative 52; Mismatches 75; Indels 55; Gaps 13;

```
Qy         14 KGAEADQIIEYLKQQVSLLKEKA---ILQATLREEKK-----LRVENAKLKKEIEEL--- 62
              | |: |::       :||   ||   :|:  | :  :     | ::  | : |  ||
Db        477 KTAKEDKVRTGTTVNISLQIVKALGILLEPFLPDASEKIWHLLNLDEVK-RWEFRELPAG 535

Qy         63 ----KQELIQAEIQNGVKQIAF------PSGTPLHANSMVS-----ENVIQSTAVTTVSS 107
                 | |::  :: :  || :        | |  |  ::      |:||:
Db        536 HKVRKPEILFKKVTD--DQIIYFILNYMAKGNPEGARILLDKYYKREDVIRV-------- 585

Qy        108 GTKEQIKGGTGDE-----KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCI 162
                ||:     |||      ::  : |: | |||| |:  :       |  | ::||||:| |
Db        586 -AKEKF----GDEAEVVLRRVYKDIKLK-EKKEGKEMYV------KFDDFAKLDLRVGKI 633

Qy        163 ITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVL 222
              |  : ||:|| ||| :||:|:   ||:|:||  :   |:: || |::: ||:| |:||:
Db        634 IEVKDHPNADKLYVVKVDLGD-EVRTLVAGLKKYYKPEELLNRYVVVVANLEPKKLRGIG 692

Qy        223 SQAMVMCASSPEKIEILAP 241
              || |:: |   |:: :| |
Db        693 SQGMLLAADDGERVALLMP 711
```

RESULT 9
D71091
methionine-tRNA ligase (EC 6.1.1.10) - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 16-Aug-2004
C;Accession: D71091
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71091
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-723 <KAW>
A;Cross-references: UNIPROT:O58721; GB:AP000004; NID:g3236131; PIDN:BAA30090.1; PID:g325
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0993
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

  Query Match 13.7%; Score 216; DB 2; Length 723;
  Best Local Similarity 29.1%; Pred. No. 8.1e-06;
  Matches 71; Conservative 53; Mismatches 72; Indels 48; Gaps 11;

```
Qy         20 QIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIEEL-------KQELIQAEIQ 72
              ||:: |     :| |  :  |: :    | ::  | | : :||       | |::  ::
Db        495 QIVKAL----GILLEPFLPDASEKIWHLLNLDEVK-KWKFKELPAGHRVRKAEILFKKVT 549

Qy         73 NGVKQIAF----------PSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE-- 120
              :  :|| :           | |        |: |   :   |  |:    ||:      |:|
Db        550 D--EQIIYFILNYMGRNNPEGA-----KMLLEKYYKREDVIKVA---KEKF----GEESK 595

Qy        121 ---KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVE 177
                 |:  : |: :  |:||: : :      |  | ::||||:| ||  : ||:|| ||:
Db        596 IILKRIYKDIKLEEGKEEKEMEYV------KFEDFAKLDLRVGKIIEVKDHPNADKLYIV 649

Qy        178 EVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIE 237
              :||:|:   ||:|:||  :   |:: |: ||:: ||:| |:||| || |:: |   | :
Db        650 KVDLGK-EVRTLVAGLKKYYKPEELLNKYVIIVANLEPKKLRGVESQGMLLAADDGENVA 708

Qy        238 ILAP 241
              :| |
Db        709 LLMP 712
```

RESULT 10
SYTWMT
methionine-tRNA ligase (EC 6.1.1.10) - Thermus aquaticus
N;Alternate names: methionyl-tRNA synthetase
C;Species: Thermus aquaticus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Dec-2002
C;Accession: A39517
R;Nureki, O.; Muramatsu, T.; Suzuki, K.; Kohda, D.; Matsuzawa, H.; Ohta, T.; Miyazawa,
J. Biol. Chem. 266, 3268-3277, 1991
A;Title: Methionyl-tRNA synthetase gene from an extreme thermophile, Thermus thermophil
esis.
A;Reference number: A39517; MUID:91131636; PMID:1993699
A;Accession: A39517
A;Molecule type: DNA
A;Residues: 1-616 <NUR>
A;Cross-references: GB:M64273; GB:J05744; NID:g155135; PIDN:AAA27510.1; PID:g155136
A;Experimental source: strain HB8, ATCC 27634
A;Note: the authors translated the codon TAC for residue 323 as Thr
C;Genetics:
A;Gene: metS
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; protein biosynthesis
F;13,19,50,300/Active site: Tyr, His, Asp, Lys #status predicted

  Query Match 13.0%; Score 205; DB 1; Length 616;
  Best Local Similarity 28.7%; Pred. No. 3.2e-05;
  Matches 71; Conservative 45; Mismatches 89; Indels 42; Gaps 10;

```
Qy         21 IIEYLKQQVSLLKEKAILQAT---LREEKKLRVENAKLKKEIEELKQELIQAEIQNGVKQ 77
              ::  ||  |:| :  | ::|    : |:|   :      ||| || :  | :  :  |::
Db        394 LVRELKFHVALEEAMAYVKALNRYINEKKPWEL----FKKEPEEARAVLYR--VVEGLR- 446

Qy         78 IAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQI------KGGTGDEKKAKEKIEKKG 131
              ||    ||   : |              | : | ||::      : |  : :   |:
Db        447 IASILLTPAMPDKM---------AELRRALGLKEEVRLEEAERWGLAEPRPIPEEAPVLF 497

Qy        132 EKKEKKQQSIAGSADSKP--------IDVSRLDLRIGCIITARKHPDADSLYVEEVDVGE 183
               ||| |         ::||          | ::::||:  :: | |||:|| | |  : :|
Db        498 PKKEAK-------VEAKPKEEAWIGIEDFAKVELRVAEVLAAEKHPNADRLLVLRLSLGN 550

Qy        184 IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAPPN 243
               ||||||:      |::  : |:|: ||||||:||: || |:: |    | :  |
Db        551 -EERTVVSGIAKWYRPEELVGKKVVLVANLKPAKLRGIESQGMILAAQEGEAL-ALVTVE 608

Qy        244 GSVPGDR 250
              | ||  |
Db        609 GEVPPGR 615
```

RESULT 11
C85968
probable tRNA synthetase ygjH [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: C85968
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <STO>
A;Cross-references: UNIPROT:Q8FDF3; GB:AE005174; NID:g12517655; PIDN:AAG58207.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ygjH
C;Superfamily: secretion chaperone CsaA

Query Match 12.6%; Score 198.5; DB 2; Length 110;
Best Local Similarity 41.7%; Pred. No. 1.1e-05;
Matches 43; Conservative 25; Mismatches 32; Indels 3; Gaps 3;

Qy 151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
Db 8 DFARLEMRVGKIVEVKRHENADKLYIVQVDVGEKTLQTVTS-LVPYYSEEELMGKTVVVL 66

Qy 211 CNLKPAKMRGVLSQAMVMCASSPEKIE-ILAPPNGSVP-GDRI 251
Db 67 CNLQKAKMRGETSECMLLCAETDDGSESVLLTPERMMPAGVRI 109

RESULT 12
D91123
probable tRNA synthetase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D91123
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D91123
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <HAY>
A;Cross-references: UNIPROT:Q8FDF3; GB:BA000007; PIDN:BAB37379.1; PID:g13363429; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3956
C;Superfamily: secretion chaperone CsaA

Query Match 12.6%; Score 198.5; DB 2; Length 110;
Best Local Similarity 41.7%; Pred. No. 1.1e-05;
Matches 43; Conservative 25; Mismatches 32; Indels 3; Gaps 3;

Qy 151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
Db 8 DFARLEMRVGKIVEVKRHENADKLYIVQVDVGEKTLQTVTS-LVPYYSEEELMGKTVVVL 66

Qy 211 CNLKPAKMRGVLSQAMVMCASSPEKIE-ILAPPNGSVP-GDRI 251
Db 67 CNLQKAKMRGETSECMLLCAETDDGSESVLLTPERMMPAGVRI 109

RESULT 13
G65095
hypothetical 12.3 kD protein in ileX-ebgR intergenic region - Escherichia coli (strain K
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: G65095
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65095
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <BLAT>
A;Cross-references: UNIPROT:P42589; GB:AE000389; GB:U00096; NID:g1789451; PIDN:AAC76109
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ygjH
C;Superfamily: secretion chaperone CsaA

Query Match 12.4%; Score 194.5; DB 2; Length 110;
Best Local Similarity 39.8%; Pred. No. 1.9e-05;
Matches 41; Conservative 27; Mismatches 32; Indels 3; Gaps 3;

Qy 151 DVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILL 210
Db 8 DFARLEMRVGKIVEVKRHENADKLYIVQVDVGQKTLQTVTS-LVPYYSEEELMGKTVVVL 66

Qy 211 CNLKPAKMRGVLSQAMVMCASSPEKIE-ILAPPNGSVP-GDRI 251
Db 67 CNLQKAKMRGETSECMLLCAETDDGSESVLLTPERMMPAGVRV 109

RESULT 14
D69431
methionine-tRNA ligase (EC 6.1.1.10) - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69431
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: D69431
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-658 <KLE>
A;Cross-references: UNIPROT:O28819; GB:AE001003; GB:AE000782; NID:g2689326; PIDN:AAB897
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: ligase

Query Match 12.1%; Score 191; DB 2; Length 658;
Best Local Similarity 27.5%; Pred. No. 0.00024;
Matches 74; Conservative 49; Mismatches 94; Indels 52; Gaps 12;

Qy 19 DQIIEYLKQQVSLLKEKAILQATLREEKK----------LRVENAK---LKKEIEE 61
Db 405 EEVLEKIRE-----TEQKILSALEEWEFKAASDAFMELATYGNVYFQNAKPWELIKTDKE 459

Qy 62 LKQELIQAEIQNGVKQIAFPSGTPLHANSMVS---------ENVIQSTA--VTTVSSGTK 110
Db 460 AARRAVASCLQLAKALIIF--AYPVMPESMERMAKAIGLDLENVRLSDAYKVDEVMKLSK 517

Qy 111 EQIK-GGTGDE--KKAKEKIEKK----GEKKEKKQQ-SIAGSADSKPIDVSRLDLRIGCI 162
Db 518 PEVPFAKVEDEVIEKLQKVMEKRFRGEGEKMEQKEEISIE--------DFQKLDIRIGRV 569

Qy 163 ITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVL 222
Db 570 LKAEKVKKSKKLIKLIIDIGD-EQRQIVSGIAEDYTPEELEGKLVVVLANLKPAKFMGVE 628

Qy 223 SQAMVMCASSPEKIEILAPPNGSVPGDRI 251

```
              |: |:: |    |  :| |     || |:
Db        629 SRGMILAAEKDGKAVLLTPEKEVEPGTRV 657


RESULT 15
E72297
methionine-tRNA ligase (EC 6.1.1.10) - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: E72297
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: E72297
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-629 <ARN>
A;Cross-references: UNIPROT:O33925; GB:AE001768; GB:AE000512; NID:g4981619; PIDN:AAD3616
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1085
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

  Query Match             11.8%;  Score 185.5;  DB 2;  Length 629;
  Best Local Similarity   26.2%;  Pred. No. 0.0005;
  Matches   50;  Conservative   46;  Mismatches   72;  Indels   23;  Gaps    4;

Qy         83 GTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDEKKAKEKIE---------- 128
              || |: ||: :   :    :  :   ::|  :  : :|| :|| :|
Db        441 GTVLY-NSLEAVFKVALMTLPVMPDTSEEVFRRVSFEEKPSKEHLENWGVLKPGSTVIHG 499

Qy        129 ----KKGEKKEKKQQSIAGSADSKPI---DVSRLDLRIGCIITARKHPDADSLYVEEVDV 181
                  || : |: |:      ||:   |   | |::|||  :: | | |::  |    :|:
Db        500 EPLFKKIDAKDFKKVVETVSAEQNAITIDDFSKVDLRIAKVLEAEKVPNSRKLLRLIIDL 559

Qy        182 GEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILAP 241
              |    | :|:|:  |   |::  :::::: ||||||: |: || |:: | | : : :|
Db        560 G-TEKRQIVAGIAEHYRPEELVGKLIVVVANLKPAKLMGIESQGMLLAAKSGDTLRLLTV 618

Qy        242 PNGSVPGDRIT 252
                 || :::
Db        619 DGEITPGAKVS 629


Search completed: January 14, 2005, 15:59:19
Job time : 51 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 15:48:39 ; Search time 190 Seconds
(without alignments)
944.826 Million cell updates/sec

Title: US-09-733-306C-4
Perfect score: 1573
Sequence: 1 MANNDAVLKRLEQKGAEADQ.....FEVKGKGVCRAQTMSNSGIK 312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1573	100.0	312	1	MCA1_HUMAN	Q12904 homo sapien
2	1362.5	86.6	359	1	MCA1_CRIGR	O54873 cricetulus
3	1343	85.4	310	1	MCA1_MOUSE	P31230 mus musculu
4	1338	85.1	319	2	Q8C2U7	Q8c2u7 mus musculu
5	1084	68.9	297	2	Q7ZWY7	Q7zwy7 xenopus lae
6	671	42.7	323	2	Q9V504	Q9v504 drosophila
7	625.5	39.8	351	2	Q7Q8S7	Q7q8s7 anopheles g
8	564	35.9	917	1	SYM_CAEEL	Q20970 caenorhabdi
9	517	32.9	804	1	SYM_ORYSA	Q9zts1 oryza sativ
10	499.5	31.8	149	2	O62542	O62542 geodia cydo
11	499	31.7	797	1	SYM_ARATH	Q9svn5 arabidopsis
12	499	31.7	797	2	AAL36365	Aal36365 arabidops
13	499	31.7	797	2	AAM14393	Aam14393 arabidops
14	453.5	28.8	529	2	Q6TGS6	Q6tgs6 brachydanio
15	453.5	28.8	529	2	AAQ97863	Aaq97863 brachydan
16	451.5	28.7	529	2	Q6DFZ7	Q6dfz7 brachydanio
17	450	28.6	528	1	SYYC_BOVIN	Q29465 bos taurus
18	450	28.6	528	1	SYYC_HUMAN	P54577 homo sapien
19	446.5	28.4	528	1	SYYC_MOUSE	Q91wq3 mus musculu
20	441.5	28.1	373	2	Q6BY38	Q6by38 debaryomyce
21	441.5	28.1	450	2	Q9P6K7	Q9p6k7 schizosacch
22	440	28.0	528	2	Q7ZX51	Q7zx51 xenopus lae
23	439	27.9	528	2	Q6DIJ1	Q6dij1 xenopus tro
24	429	27.3	376	1	G4P1_YEAST	P46672 saccharomyc
25	429	27.3	376	2	AAS56824	Aas56824 saccharom
26	426	27.1	376	2	Q6FNN0	Q6fnn0 candida gla
27	423.5	26.9	371	2	Q754V3	Q754v3 ashbya goss
28	423.5	26.9	371	2	AAS53340	Aas53340 ashbya go
29	415	26.4	372	2	Q6CKI6	Q6cki6 kluyveromyc
30	402	25.6	409	2	Q8RUP8	Q8rup8 oryza sativ
31	401.5	25.5	394	2	Q6C763	Q6c763 yarrowia li
32	400.5	25.5	389	2	Q93VB0	Q93vb0 arabidopsis
33	400.5	25.5	440	2	Q7XJM9	Q7xjm9 arabidopsis
34	395	25.1	525	2	Q9VV60	Q9vv60 drosophila
35	391	24.9	252	2	Q7XVQ8	Q7xvq8 oryza sativ
36	388	24.7	78	2	Q9EPV3	Q9epv3 rattus norv
37	378.5	24.1	419	2	Q7SAC3	Q7sac3 neurospora
38	343	21.8	273	2	Q9M1X8	Q9m1x8 arabidopsis
39	337.5	21.5	441	2	O97487	O97487 euplotes oc
40	318.5	20.2	542	2	Q7QD89	Q7qd89 anopheles g
41	292.5	18.6	402	2	Q8IL48	Q8il48 plasmodium
42	264	16.8	403	2	Q7RKA7	Q7rka7 plasmodium
43	231.5	14.7	724	1	SYM_PYRFU	Q8u221 pyrococcus
44	227	14.4	197	2	Q6BTQ8	Q6btq8 debaryomyce
45	223.5	14.2	722	1	SYM_PYRAB	Q9v011 pyrococcus

ALIGNMENTS

RESULT 1
MCA1_HUMAN
ID MCA1_HUMAN STANDARD; PRT; 312 AA.
AC Q12904; Q96CQ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Multisynthetase complex auxiliary component p43 [Contains:
DE Endothelial-monocyte activating polypeptide II (EMAP-II) (Small
DE inducible cytokine subfamily E member 1)].
GN Name=SCYE1; Synonyms=EMAP2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95014290; PubMed=7929199;
RA Kao J., Houck K., Fan Y., Haehnel I., Libutti S.K., Kayton M.L.,
RA Grikscheit T., Chabot J., Nowygrod R., Greenberg S., Kuang W.J.,
RA Leung D.W., Hayward J.R., Kisiel W., Heath M., Brett J., Stern D.M.;
RT "Characterization of a novel tumor-derived cytokine. Endothelial-
RT monocyte activating polypeptide II.";
RL J. Biol. Chem. 269:25106-25119(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U10117; AAA62202.1; -.
DR   EMBL; BC014051; AAH14051.1; -.
DR   PDB; 1E7Z; X-ray; A=147-312.
DR   PDB; 1EUJ; X-ray; A/B=147-312.
DR   PDB; 1FL0; X-ray; A=150-312.
DR   Genew; HGNC:10648; SCYE1.
DR   MIM; 603605; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005125; F:cytokine activity; TAS.
DR   GO; GO:0000049; F:tRNA binding; TAS.
DR   GO; GO:0006935; P:chemotaxis; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   GO; GO:0006418; P:tRNA aminoacylation for protein translation; TAS.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   3D-structure; Cytokine; Protein biosynthesis; RNA-binding;
KW   tRNA-binding.
FT   PROPEP        1    146
FT   CHAIN       147    312       Endothelial-monocyte activating
FT                                polypeptide II.
FT   DOMAIN      151    252       tRNA-binding.
FT   CONFLICT     79     79       A -> P (in Ref. 2).
FT   HELIX       152    154
FT   STRAND      157    168
FT   TURN        169    170
FT   STRAND      171    180
FT   STRAND      187    191
FT   TURN        194    196
FT   HELIX       199    202
FT   TURN        203    204
FT   STRAND      205    210
FT   STRAND      216    217
FT   STRAND      222    223
FT   STRAND      226    227
FT   STRAND      229    232
FT   TURN        233    234
FT   STRAND      235    238
FT   STRAND      240    240
FT   TURN        243    244
FT   TURN        247    248
FT   STRAND      250    250
FT   TURN        254    255
FT   TURN        266    268
FT   TURN        270    270
FT   HELIX       271    274
FT   HELIX       275    277
FT   STRAND      278    280
FT   TURN        282    283
FT   STRAND      285    288
FT   TURN        289    290
FT   STRAND      291    292
FT   STRAND      294    295
FT   TURN        296    298
FT   STRAND      299    300
FT   STRAND      302    302
FT   STRAND      310    312
SQ   SEQUENCE   312 AA;  34326 MW;  946310A0216F7587 CRC64;
```

Query Match 100.0%; Score 1573; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 8.5e-86;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300

Qy        301 CRAQTMSNSGIK 312
              ||||||||||||
Db        301 CRAQTMSNSGIK 312
```

```
RESULT 2
MCA1_CRIGR
ID   MCA1_CRIGR     STANDARD;      PRT;   359 AA.
AC   O54873;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Multisynthetase complex auxiliary component p43 [Contains:
DE   Endothelial-monocyte activating polypeptide II (EMAP-II) (Small
DE   inducible cytokine subfamily E member 1)].
GN   Name=SCYE1;
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98070438; PubMed=9405472;
RA   Quevillon S., Agou F., Robinson J.-C., Mirande M.;
RT   "The p43 component of the mammalian multi-synthetase complex is likely
RT   to be the precursor of the endothelial monocyte-activating polypeptide
RT   II cytokine.";
RL   J. Biol. Chem. 272:32573-32579(1997).
CC   -!- SUBUNIT: Component of the multisynthetase complex which is
CC       comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC       monospecific isoleucyl, leucyl, glutaminyl, methionyl, lysyl,
CC       arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC       proteins, p18, p48 and p43.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF021800; AAB95207.1; -.
DR   HSSP; Q12904; 1E7Z.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Cytokine; Protein biosynthesis; RNA-binding; tRNA-binding.
FT   DOMAIN      198    299       tRNA-binding.
SQ   SEQUENCE   359 AA;  39601 MW;  4D868D1B65D72C23 CRC64;
```

```
  Query Match             86.6%;  Score 1362.5;  DB 1;  Length 359;
  Best Local Similarity   86.0%;  Pred. No. 3.2e-73;
  Matches  270;  Conservative   19;  Mismatches   22;  Indels    3;  Gaps    2;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:||||||:|||||||||||||||||||||||
Db         47 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAVLQATLREEKKLRVENAKLKKEIE 106

Qy         61 ELKQELIQAEIQNGVKQIAFP--SGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTG 118
              |||||||||||||||||  |  | ||: |:| || :|||||:|:|:|   ||  ||| |
Db        107 ELKQELIQAEIQNGVKQIPVPVQSDTPVQASSAVSTSVIQSTSVSTISCSIKEHSKGG-G 165

Qy        119 DEKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEE 178
              :||| ||| :||||||||| || | ||||||:|||||||||| |:| :||||||||||||
Db        166 EEKKVKEKTDKKGEKKEKKLQSAAPSADSKPVDVSRLDLRIGRIVTVKKHPDADSLYVEE 225

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEI 238
              ||||| |||||:|||||||||:|||||||:||||||||||:|||||||||||||||:||
Db        226 VDVGEAAPRTVISGLVNHVPLDQMQNRMVVLLCNLKPAKMRGILSQAMVMCASSPEKVEI 285

Qy        239 LAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGK 298
              ||||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db        286 LAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGSPFEVKGK 345

Qy        299 GVCRAQTMSNSGIK 312
              ||||||||:|||||
Db        346 GVCRAQTMANSGIK 359


RESULT 3
MCA1_MOUSE
ID   MCA1_MOUSE     STANDARD;      PRT;   310 AA.
AC   P31230; Q60659;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Multisynthetase complex auxiliary component p43 [Contains:
DE   Endothelial-monocyte activating polypeptide II (EMAP-II) (Small
DE   inducible cytokine subfamily E member 1)].
GN   Name=Scye1; Synonyms=Emap2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95014290; PubMed=7929199;
RA   Kao J., Houck K., Fan Y., Haehnel I., Libutti S.K., Kayton M.L.,
RA   Grikscheit T., Chabot J., Nowygrod R., Greenberg S., Kuang W.J.,
RA   Leung D.W., Hayward J.R., Kisiel W., Heath M., Brett J., Stern D.M.;
RT   "Characterization of a novel tumor-derived cytokine. Endothelial-
RT   monocyte activating polypeptide II.";
RL   J. Biol. Chem. 269:25106-25119(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 145-164 FROM N.A.
RX   MEDLINE=93015897; PubMed=1400342;
RA   Kao J., Ryan J., Brett G., Chen J., Shen H., Fan Y.-G., Godman G.,
RA   Familletti P.C., Wang F., Pan Y.-C.E., Stern D., Clauss M.;
RT   "Endothelial monocyte-activating polypeptide II. A novel tumor-derived
RT   polypeptide that activates host-response mechanisms.";
RL   J. Biol. Chem. 267:20239-20247(1992).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=94193665; PubMed=7545917;
RA   Kao J., Fan Y., Haehnel I., Brett J., Greenberg S., Clauss M.,
RA   Kayton M., Houck K., Kisiel W., Seljelid R., Burnier J., Stern D.;
RT   "A peptide derived from the amino terminus of endothelial-monocyte-
RT   activating polypeptide II modulates mononuclear and polymorphonuclear
RT   leukocyte functions, defines an apparently novel cellular interaction
RT   site, and induces an acute inflammatory response.";
RL   J. Biol. Chem. 269:9774-9782(1994).
CC   -!- FUNCTION: Alters endothelial and monocyte functions, induces the
CC       migration of monocytes and granulocytes, and induces an
CC       inflammatory response in the mouse footpad model. EMAP II elicits
CC       a phlogogenic response and, potentially, augments the effects of
CC       the other tumor-derived cytokines.
CC   -!- SUBUNIT: Monomer.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10118; AAA62203.1; -.
DR   EMBL; BC002054; AAH02054.1; -.
DR   PIR; A55053; A55053.
DR   HSSP; Q12904; 1EUJ.
DR   MGD; MGI:102774; Scye1.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Cytokine; Protein biosynthesis; RNA-binding; tRNA-binding.
FT   PROPEP          1    144
FT   CHAIN         145    310       Endothelial-monocyte activating
FT                                   polypeptide II.
FT   DOMAIN        149    250       tRNA-binding.
SQ   SEQUENCE   310 AA;  33997 MW;  A2F8FF52A33D03A0 CRC64;

  Query Match             85.4%;  Score 1343;  DB 1;  Length 310;
  Best Local Similarity   85.9%;  Pred. No. 4e-72;
  Matches  269;  Conservative   13;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||:||||||:|||||||||||||||||||||||
Db          1 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         61 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASPATKEQIK--AGE 118

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||||||:||:||||||||||||
Db        119 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRIGCIVTAKKHPDADSLYVEEV 177

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||||||:|||||||||||||||||||||||||||:|||
```

```
Db        178 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 237

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        238 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 297

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        298 VCRAQTMANSGIK 310


RESULT 4
Q8C2U7
ID   Q8C2U7      PRELIMINARY;      PRT;   319 AA.
AC   Q8C2U7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE   length enriched library, clone:E430001B22 product:endothelial monocyte
DE   activating polypeptide 2, full insert sequence.
GN   Name=Scye1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK087932; BAC40045.1; -.
DR   HSSP; Q12904; 1EUJ.
DR   MGD; MGI:102774; Scye1.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0000049; F:tRNA binding; IEA.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PROSITE; PS50886; TRBD; 1.
SQ   SEQUENCE   319 AA;  35197 MW;  B977E8FE742BEEC6 CRC64;

  Query Match             85.1%;  Score 1338;  DB 2;  Length 319;
  Best Local Similarity   85.3%;  Pred. No. 8.1e-72;
  Matches  267;  Conservative   15;  Mismatches   27;  Indels    4;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || |||||||||||||||||||||||||||:|||||||||||:|||||||||||||:|||
Db         10 MATNDAVLKRLEQKGAEADQIIEYLKQQVALLKEKAILQATMREEKKLRVENAKLKREIE 69

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAV-TTVSSGTKEQIKGGTGD 119
              ||||||| ||| |||:|:     |||  |   ||:|:|| :| || |  ||||||   |:
Db         70 ELKQELILAEIHNGVEQVRVRLSTPLQTNCTASESVVQSPSVATTASLATKEQIK--AGE 127

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
              ||| ||| |||||||| |||| | | |||||| ||||||:|||:||:|||||||||||||
Db        128 EKKVKEKTEKKGEKKE-KQQSAAASTDSKPIDASRLDLRVGCIVTAKKHPDADSLYVEEV 186

Qy        180 DVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEIL 239
              |||| |||||||||||||||||||||:|||||||||||||||||||||||||:|||
Db        187 DVGEAAPRTVVSGLVNHVPLEQMQNRMVVLLCNLKPAKMRGVLSQAMVMCASSPEKVEIL 246

Qy        240 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKG 299
              ||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db        247 APPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNAECVATYKGAPFEVKGKG 306

Qy        300 VCRAQTMSNSGIK 312
              |||||||:|||||
Db        307 VCRAQTMANSGIK 319


RESULT 5
Q7ZWY7
ID   Q7ZWY7      PRELIMINARY;      PRT;   297 AA.
AC   Q7ZWY7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Scye1-prov protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC046580; AAH46580.1; -.
DR   HSSP; Q12904; 1FL0.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0000049; F:tRNA binding; IEA.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
SQ   SEQUENCE   297 AA;  32510 MW;  C9D40417020CAA11 CRC64;

  Query Match             68.9%;  Score 1084;  DB 2;  Length 297;
  Best Local Similarity   71.9%;  Pred. No. 9.4e-57;
  Matches  217;  Conservative   30;  Mismatches   47;  Indels    8;  Gaps    3;

Qy          1 MANNDAVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIE 60
              || ::  || ||:|: |||||||||||||:||||||||::|||||||||||||||||||
Db          1 MATSNPVLNRLDQRAAEADQIIEYLKQQVALLKEKAILQASVREEKKLRVENAKLKKEIE 60

Qy         61 ELKQELIQAEIQNGVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTGDE 120
               ||:::: ||:||||||: |: |   |:| ||  |          ||    :    :|:|
Db         61 VLKEQIVTTEIKNGVKQISIPTST--SADSSVSAPVSAPQPAPVKSSPPAPK----SGEE 114

Qy        121 KKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVD 180
              || ||| ||||||||||    |   : | :|||||||:||||||:||||||||||||||||
Db        115 KKKKEKAEKKGEKKEKKPP--ASEDELKAVDVSRLDLRVGCIITAKKHPDADSLYVEEVD 172

Qy        181 VGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEILA 240
              |||  ||||||||| |:|||||||| :|||||||||||:|||||||||||||:|||
Db        173 VGEATPRTVVSGLVKHIPLEQMQNRMAVLLCNLKPAKMRGILSQAMVMCASSPEKVEILD 232

Qy        241 PPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGV 300
              ||:|:|||||||||  |||||||||||||| |||||||| |||:||||||| ||||:|||
Db        233 PPSGAVPGDRITFQGFPGEPDKELNPKKKTWEQIQPDLLTNDKCVATYKGAPFEVQGKGC 292

Qy        301 CR 302
              ||
Db        293 AR 294


RESULT 6
Q9V504
ID   Q9V504      PRELIMINARY;      PRT;   323 AA.
AC   Q9V504;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG8235-PA.
GN   ORFNames=CG8235;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
```

```
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FLYBASE;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FLYBASE;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003835; AAF59019.2; -.
DR   HSSP; Q12904; 1EUJ.
DR   IntAct; Q9V504; -.
DR   FlyBase; FBgn0033351; CG8235.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0000049; F:tRNA binding; IEA.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
SQ   SEQUENCE   323 AA;  34401 MW;  6EEB6E30115F4E84 CRC64;

  Query Match             42.7%;  Score 671;  DB 2;  Length 323;
  Best Local Similarity   44.9%;  Pred. No. 3.7e-32;
  Matches  141;  Conservative   57;  Mismatches   88;  Indels   28;  Gaps    5;

Qy          6 AVLKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIEELKQE 65
              | |:::      |: :|  :: ::| :::    |   |::::|  ||| | ||:|    :
Db         31 ADLQQIASNNERAEALINSIEAEISGIQQ----QLVERQKQELIKENAALAKEVEAALAQ 86

Qy         66 LIQAEIQNGVKQIAFP-------SGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGTG 118
              |:| |::|| ||| |       |  |:    ::     :||    :
Db         87 LVQLELRNGKKQIPVPGARGFCTSAAPV----VMPAEAGPATAAPAAPA---------- 131

Qy        119 DEKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEE 178
               | |||  ||| ::|:   :  | :|   |:|| |||||:| |:   :|||||||:|:
Db        132 -PKPAKEPKEKKSKEKKPAAEKPA-AAPEAPVDVGRLDLRVGKIVEVGRHPDADSLYLEK 189

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEI 238
              :| || |||||||||| ||||:||||:|:::|||||||||| |:|||||:|||:|:
Db        190 IDCGEAAPRTVVSGLVKFVPLEEMQNRLVVVMCNLKPAKMRGVTSEAMVMCASTPEKVEV 249

Qy        239 LAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGK 298
              |:|| |:|||| :  : :| :|| :|||||||:|   ||| || | || |||    | ||
Db        250 LSPPAGAVPGDLVHCEGYPRQPDAQLNPKKKIFESCAPDLKTNGELVACYKGAALHVPGK 309

Qy        299 GVCRAQTMSNSGIK 312
              |   |||: |  :|
Db        310 GNVVAQTLKNVNVK 323


RESULT 7
Q7Q8S7
ID   Q7Q8S7      PRELIMINARY;      PRT;   351 AA.
AC   Q7Q8S7;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP11872 (Fragment).
GN   Name=agCG47516; ORFNames=ENSANGG00000014444;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008933; EAA09959.1; -.
DR   GO; GO:0000049; F:tRNA binding; IEA.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PROSITE; PS50886; TRBD; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   351 AA;  38153 MW;  F3F1E2D23B1E3DFE CRC64;

  Query Match             39.8%;  Score 625.5;  DB 2;  Length 351;
  Best Local Similarity   42.7%;  Pred. No. 2e-29;
  Matches  134;  Conservative   53;  Mismatches   94;  Indels   33;  Gaps    5;

Qy          8 LKRLEQKGAEADQIIEYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEIEELKQELI 67
              |:|:      |::::| |||:                        ||  |:::::   ::|:
Db         62 LQRILSNNKAAEELLESLKQE--------------------TENETLRRQVDSCLKQLV 100

Qy         68 QAEIQNGVKQIAFPSGTPLHANS----MVSENVIQSTAVTTVSSGTKEQ----IKGGTGD 119
                |: :|  ||  |:   : | |     :  | | |  |          |||     :   |
Db        101 SLEVAHGKVQIPVPTPDEVRAGSGPATVKMEPVAQPPA---AKESPKEQPPAKAQQKQSD 157

Qy        120 EKKAKEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEV 179
               |  |||  || :    : :  | : :  |||| |||:|:| |:   :||||||||||::
Db        158 SKPPKEKKPKKEKPAGGEGKPAAPAVEEPPIDVGRLDMRVGRIVEVSRHPDADSLYVEKI 217

Qy        180 DVGE-IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPEKIEI 238
               | ||   ||||:|||| :||:|:|||||: ||||||||||:||:|||||||:|:::||
Db        218 DCGEPNGPRTVISGLVKYVPIEEMQNRMVVALCNLKPAKMRGILSEAMVMCASTPDRVEI 277

Qy        239 LAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGK 298
              ||||  :|||| :  : :|  ||  :||||||:| : ||| || | || ||    | | ||
Db        278 LAPPADAVPGDLVHVEGYPRVPDSVMNPKKKIFETVAPDLKTNGELVACYKEGSFVVPGK 337

Qy        299 GVCRAQTMSNSGIK 312
              |  :|||: |  :|
Db        338 GPVKAQTLKNVQVK 351


RESULT 8
SYM_CAEEL
ID   SYM_CAEEL      STANDARD;      PRT;   917 AA.
AC   Q20970;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE   (MetRS).
GN   Name=mrs-1; ORFNames=F58B3.5;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Harris B.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
```

```
CC       diphosphate + L-methionyl-tRNA(Met).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z73427; CAA97803.1; -.
DR   PIR; T22898; T22898.
DR   HSSP; Q12904; 1FL0.
DR   WormPep; F58B3.5; CE06007.
DR   InterPro; IPR008224; MetRS_dimerising.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002300; tRNA-synt_1a.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR002304; tRNA-synt_met.
DR   InterPro; IPR009080; tRNAsyn_1a_bind.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF00133; tRNA-synt_1; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR   PRINTS; PR01041; TRNASYNTHMET.
DR   TIGRFAMs; TIGR00398; metG; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Ligase; Protein biosynthesis;
KW   RNA-binding; tRNA-binding.
FT   SITE         44     54       "HIGH" region.
FT   SITE        367    371       "KMSKS" region.
FT   BINDING     370    370       ATP (By similarity).
FT   DOMAIN      756    857       tRNA-binding.
SQ   SEQUENCE   917 AA;  101713 MW;  3B42AABA314FBDF0 CRC64;

  Query Match             35.9%;  Score 564;  DB 1;  Length 917;
  Best Local Similarity   42.3%;  Pred. No. 2.7e-25;
  Matches  135;  Conservative   52;  Mismatches  100;  Indels   32;  Gaps    8;

Qy          1 MANNDAVLKRLEQKGAEADQII-EYLKQQVSLLKEKAILQATLREEKKLRVENAKLKKEI 59
              ||:  | :: |||      |:| : ||:     : ||:   |  : ::|  || :|  ::
Db        624 MASTAAFVE-LEQGAKVISQLIAQNLKK---FDQAKALF--TRNQLQRLDGENKQLTIDV 677

Qy         60 EELKQELIQAEIQNGVKQIAFP--SGTPLHANSMVSENVIQSTAVTTVSSGTKEQIKGGT 117
              : |: :||: |   |:||:  |  | ||              |  :| :||   :
Db        678 KTLQHQLIELETAAGIKQVPKPVVSCTP-------------TPTSTPASGIITE----- 718

Qy        118 GDEKKAKEKIEKKGEKKEKKQQSIAGSADSKP----IDVSRLDLRIGCIITARKHPDADS 173
                :|:|          || |:|:   | | : |    ||| |||:|:| ||   ||||||:
Db        719 APKKEAPSTPAPSEPKKAKEQKKGKGGAAAAPVDDTIDVGRLDMRVGRIIKCEKHPDADA 778

Qy        174 LYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSP 233
              ||||::|||| |||||||||| ||||:|||||:|::|||||||||||| |:||||||||
Db        779 LYVEQIDVGESAPRTVVSGLVRHVPLDQMQNRLVVVLCNLKPAKMRGVESRAMVMCASSP 838

Qy        234 EKIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPF 293
              :|:||:  |  | ||  :    :   ||::|||||||| :  ||  : |  | :|| |
Db        839 DKVEIMEVPADSKPGTPVVCPPYTHRPDEQLNPKKKIWETVAEDLKVSAEGFAEWKGQPL 898

Qy        294 EVKGKGVCRAQTMSNSGIK 312
               :  :    | |:   :|
Db        899 LIGSESKMTAPTLRGVHVK 917


RESULT 9
SYM_ORYSA
ID   SYM_ORYSA      STANDARD;      PRT;   804 AA.
AC   Q9ZTS1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA
DE   ligase) (MetRS).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Deniziak M., Mirande M.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC       diphosphate + L-methionyl-tRNA(Met).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF040700; AAC99620.1; -.
DR   HSSP; Q12904; 1FL0.
DR   Gramene; Q9ZTS1; -.
DR   InterPro; IPR008224; MetRS_dimerising.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002300; tRNA-synt_1a.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR002304; tRNA-synt_met.
DR   InterPro; IPR009080; tRNAsyn_1a_bind.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF00133; tRNA-synt_1; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR   PRINTS; PR01041; TRNASYNTHMET.
DR   TIGRFAMs; TIGR00398; metG; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Ligase; Protein biosynthesis;
KW   RNA-binding; tRNA-binding.
FT   DOMAIN        4     11       Poly-Pro.
FT   SITE         28     38       "HIGH" region.
FT   SITE        350    354       "KMSKS" region.
FT   BINDING     353    353       ATP (By similarity).
FT   DOMAIN      642    745       tRNA-binding.
SQ   SEQUENCE   804 AA;  89716 MW;  A87E8ABBD419D440 CRC64;

  Query Match             32.9%;  Score 517;  DB 1;  Length 804;
  Best Local Similarity   38.1%;  Pred. No. 1.4e-22;
  Matches  117;  Conservative   53;  Mismatches   79;  Indels   58;  Gaps    9;

Qy         35 KAILQATLREEKKL-----RVENAKLKKE---------------IEELKQELIQAEIQN 73
              : :||  :  |: |     : | || |:                   :||| | :::
Db        527 EVLLQLNMTPEESLSFCDDKGEIAKAKRPWDFVSAGHKIGKPSPLFKELKDEEVES---- 582

Qy         74 GVKQIAFPSGTPLHANSMVSENVIQSTAVTTVSSGTK--EQIKGGTGDEKKAKEKIEKKG 131
                           |       :|:     :   |  :::||          |:   |
Db        583 -------------FRNKFAGSQAERSSKAQADAEAKKVADKLKG---------TKLSDGG 620

Qy        132 EKKEKKQQSIAGSADSK----PIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGEIAPR 187
              :|||:|:||  | : ||     : |::||:|:| |  |:|||||||||||:|||| |||
Db        621 QKKEQKKQS--GGSKSKNAEVDVTVAKLDIRVGLIRKAQKHPDADSLYVEEIDVGEEAPR 678

Qy        188 TVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAPPNGS 245
              |||||||  :|||:|||| | :|||||  |||: | |||: ||: :  |:|:: ||  :
Db        679 TVVSGLVKFIPLEEMQNRKVCVLCNLKPVAMRGIKSHAMVLAASNEDHTKVELVEPPESA 738
```

```
Qy        246 VPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVCRAQT 305
              |:|:||  : |||:  || | | ||::  |||:| | || || |||     |||: ::
Db        739 AVGERVTFAGYSGEPEASLNAKSKTWEKLSADLHSNGELVACYKDVPFTTSA-GVCKVKS 797

Qy        306 MSNSGIK 312
              :::  |:
Db        798 IASGEIR 804


RESULT 10
O62542
ID   O62542      PRELIMINARY;      PRT;   149 AA.
AC   O62542;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Endothelial-monocyte-activating polypeptide related protein.
GN   Name=EMAPR1;
OS   Geodia cydonium (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Astrophorida; Geodiidae; Geodia.
OX   NCBI_TaxID=6047;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98184040; PubMed=9523439;
RA   Pahler S., Krasko A., Schuetze J., Mueller I.M., Mueller W.E.G.;
RT   "Isolation and characterisation of the cDNA, encoding a potential
RT   morphogen from the marine sponge Geodia cydonium that is conserved in
RT   higher Metazones.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 265:421-425(1998).
DR   EMBL; Y14947; CAA75164.1; -.
DR   HSSP; Q12904; 1FL0.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0000049; F:tRNA binding; IEA.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PROSITE; PS50886; TRBD; 1.
SQ   SEQUENCE   149 AA;  16499 MW;  61272889BE17C7EF CRC64;

  Query Match             31.8%;  Score 499.5;  DB 2;  Length 149;
  Best Local Similarity   62.2%;  Pred. No. 2.4e-22;
  Matches   92;  Conservative   24;  Mismatches   31;  Indels    1;  Gaps    1;

Qy        157 LRIGCIITARKHPDADSLYVEEVDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPA 216
              :||| | :  :|||||:||||::||||  |||| |||| || :| | ||:|::||||||
Db          1 MRIGRITSVERHPDADTLYVEQIDVGEEKPRTVCSGLVTHVAIETMNNRLVVVLCNLKPV 60

Qy        217 KMRGVLSQAMVMCASSPEKIEILAPPNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQP 276
              ||||| |:|||||||||| |||| ||:  |||||:||  : | || :||||||::| :||
Db         61 KMRGVTSEAMVMCASSPENIEILDPPDSCVPGDRVTFTGYTGSPDTQLNPKKKVFETVQP 120

Qy        277 DLHTNDECVATYKGVPFEVKGKGVCRAQ 304
              |   |:| ||||:|:| : : : || ||
Db        121 DFLVNEEGVATYRGIPLQWRVR-VCAAQ 147


RESULT 11
SYM_ARATH
ID   SYM_ARATH      STANDARD;      PRT;   797 AA.
AC   Q9SVN5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Probable methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA
DE   ligase) (MetRS).
GN   OrderedLocusNames=At4g13780; ORFNames=F18A5.170;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC       diphosphate + L-methionyl-tRNA(Met).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL035528; CAB36842.1; -.
DR   EMBL; AL161537; CAB78420.1; -.
DR   PIR; T05247; T05247.
DR   HSSP; Q12904; 1E7Z.
DR   InterPro; IPR008224; MetRS_dimerising.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002300; tRNA-synt_1a.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR002304; tRNA-synt_met.
DR   InterPro; IPR009080; tRNAsyn_1a_bind.
```

```
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF00133; tRNA-synt_1; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR   PRINTS; PR01041; TRNASYNTHMET.
DR   TIGRFAMs; TIGR00398; metG; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Ligase; Protein biosynthesis;
KW   RNA-binding; tRNA-binding.
FT   SITE         26     36       "HIGH" region.
FT   SITE        348    352       "KMSKS" region.
FT   BINDING     351    351       ATP (By similarity).
FT   DOMAIN      635    738       tRNA-binding.
SQ   SEQUENCE   797 AA;  89853 MW;  B0F523C2E0C1017F CRC64;

  Query Match             31.7%;  Score 499;  DB 1;  Length 797;
  Best Local Similarity   52.9%;  Pred. No. 1.7e-21;
  Matches  101;  Conservative   31;  Mismatches   51;  Indels    8;  Gaps    3;

Qy        124 KEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGE 183
              |:|   ||  | | |      |  : | ::|||:|:| |: | ||| ||:|||||:|||
Db        613 KQKASSKGGGKPKPQ-----PAADREITMARLDIRVGKIVKAEKHPKADALYVEEIDVGG 667

Qy        184 IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAP 241
              |||||||| ::|||:|||||| :|||||||||| ::|||||: ||| :  |:|:: |
Db        668 GEIRTVVSGLVKYIPLEEMQNRMVCVLCNLKPAKMRDIVSQAMVLAASSSDGSKVELVEP 727

Qy        242 PNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVC 301
              |  :  |:|:||  | ||||  ||||||:|| :  ||:| :  || || |||     |||
Db        728 PKTANIGERVTFPGFEGEPDDVLNPKKKVWETLLVDLNTKENLVACYKDVPFTTSA-GVC 786

Qy        302 RAQTMSNSGIK 312
              :  ::||  |:
Db        787 KVSSISNGTIR 797


RESULT 12
AAL36365
ID   AAL36365    PRELIMINARY;      PRT;   797 AA.
AC   AAL36365;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative methionyl-tRNA synthetase.
GN   AT4G13780.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RT   "Arabidopsis Full Length cDNA Clones.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY064009; AAL36365.1; -.
KW   Aminoacyl-tRNA synthetase.
SQ   SEQUENCE   797 AA;  89853 MW;  B0F523C2E0C1017F CRC64;

  Query Match             31.7%;  Score 499;  DB 2;  Length 797;
  Best Local Similarity   52.9%;  Pred. No. 1.7e-21;
  Matches  101;  Conservative   31;  Mismatches   51;  Indels    8;  Gaps    3;

Qy        124 KEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGE 183
              |:|   ||  | | |      |  : | ::|||:|:| |: | ||| ||:|||||:|||
Db        613 KQKASSKGGGKPKPQ-----PAADREITMARLDIRVGKIVKAEKHPKADALYVEEIDVGG 667

Qy        184 IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAP 241
              |||||||| ::|||:|||||| :|||||||||| ::|||||: ||| :  |:|:: |
Db        668 GEIRTVVSGLVKYIPLEEMQNRMVCVLCNLKPAKMRDIVSQAMVLAASSSDGSKVELVEP 727

Qy        242 PNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVC 301
              |  :  |:|:||  | ||||  ||||||:|| :  ||:| :  || || |||     |||
Db        728 PKTANIGERVTFPGFEGEPDDVLNPKKKVWETLLVDLNTKENLVACYKDVPFTTSA-GVC 786

Qy        302 RAQTMSNSGIK 312
              :  ::||  |:
Db        787 KVSSISNGTIR 797


RESULT 13
AAM14393
ID   AAM14393    PRELIMINARY;      PRT;   797 AA.
AC   AAM14393;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative methionyl-tRNA synthetase.
GN   AT4G13780.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RT   "Arabidopsis Open Reading Frame (ORF) Clones.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY091454; AAM14393.1; -.
KW   Aminoacyl-tRNA synthetase.
SQ   SEQUENCE   797 AA;  89853 MW;  B0F523C2E0C1017F CRC64;

  Query Match             31.7%;  Score 499;  DB 2;  Length 797;
  Best Local Similarity   52.9%;  Pred. No. 1.7e-21;
  Matches  101;  Conservative   31;  Mismatches   51;  Indels    8;  Gaps    3;

Qy        124 KEKIEKKGEKKEKKQQSIAGSADSKPIDVSRLDLRIGCIITARKHPDADSLYVEEVDVGE 183
              |:|   ||  | | |      |  : | ::|||:|:| |: | ||| ||:|||||:|||
Db        613 KQKASSKGGGKPKPQ-----PAADREITMARLDIRVGKIVKAEKHPKADALYVEEIDVGG 667

Qy        184 IAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCASSPE--KIEILAP 241
              |||||||| ::|||:|||||| :|||||||||| ::|||||: ||| :  |:|:: |
Db        668 GEIRTVVSGLVKYIPLEEMQNRMVCVLCNLKPAKMRDIVSQAMVLAASSSDGSKVELVEP 727

Qy        242 PNGSVPGDRITFDAFPGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFEVKGKGVC 301
              |  :  |:|:||  | ||||  ||||||:|| :  ||:| :  || || |||     |||
Db        728 PKTANIGERVTFPGFEGEPDDVLNPKKKVWETLLVDLNTKENLVACYKDVPFTTSA-GVC 786

Qy        302 RAQTMSNSGIK 312
              :  ::||  |:
Db        787 KVSSISNGTIR 797


RESULT 14
Q6TGS6
ID   Q6TGS6      PRELIMINARY;      PRT;   529 AA.
AC   Q6TGS6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Tyrosyl-tRNA synthetase.
```

```
GN   Name=YARS;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney marrow;
RA   Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,
RA   Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zon L.I.,
RA   Kanki J.P., Look A.T., Chen Z.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY398430; AAQ97863.1; -.
DR   GO; GO:0004812; F:tRNA ligase activity; IEA.
DR   InterPro; IPR008232; EMAPII.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR002305; tRNA-synt_1b.
DR   InterPro; IPR002547; tRNA_bind.
DR   InterPro; IPR002307; Tyr_tRNA-synt_1b.
DR   Pfam; PF00579; tRNA-synt_1b; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF005381; EMAPII; 1.
DR   PRINTS; PR01040; TRNASYNTHTYR.
DR   TIGRFAMs; TIGR00234; tyrS; 1.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase.
SQ   SEQUENCE   529 AA;  59505 MW;  C3314003813B9336 CRC64;

  Query Match             28.8%;  Score 453.5;  DB 2;  Length 529;
  Best Local Similarity   39.8%;  Pred. No. 5.4e-19;
  Matches  123;  Conservative   43;  Mismatches   78;  Indels   65;  Gaps   14;

Qy         44 EEKKLRV--ENAKLKKEIEELKQELIQAEIQNG----VKQIAFPSGTPLHANSMVS---- 93
              || |: :  :|  :||::::   |    |  ||    || : |    |||:  ::
Db        228 EESKIDLLDKNQDVKKKLKKAFCEPGNVE-NNGVLSFVKHVLF----PLHSEFVIKRDPK 282

Qy         94 ----------ENVIQSTAVTTVSSGTKEQIKGGTGDEK--------KAKEKIEKKGEKKE 135
                       | | :  |         |||    || |          |  : | || |  |
Db        283 FGGDKVYTDFEEVEKDFAA--------EQIH--PGDLKASVELALNKLLDPIRKKFESPE 332

Qy        136 -KKQQSIA--------GSADSKP--------IDVSRLDLRIGCIITARKHPDADSLYVEE 178
               ||  | |         | |   |         :  ||||:|:| :|:  |||||||||:|:
Db        333 LKKLTSSAYPEPSKNKGGAKGNPKQTTDDDEVIPSRLDIRVGKVISVEKHPDADSLYLEK 392

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCAS---SPEK 235
              :|||| |||||||| ::  ||:|:|:|:|||||| ||||: |||||:|||    | |
Db        393 IDVGEEQPRTVVSGLVAYITEEQLQDRLVVLLCNLKPQKMRGIESQAMVLCASIEGEPRK 452

Qy        236 IEILAPPNGSVPGDRITFDAF-PGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFE 294
              :| | || ||  |||:  : :  |:|| || ||||::|::| ||  : | || :|
Db        453 VEPLDPPEGSAAGDRVYVEGYESGKPDDELKPKKKVFEKLQVDLKISGEFVAQWKEQNLM 512

Qy        295 VK-GKGVCR 302
               | |:  |:
Db        513 TKLGRITCK 521


RESULT 15
AAQ97863
ID   AAQ97863    PRELIMINARY;      PRT;   529 AA.
AC   AAQ97863;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Tyrosyl-tRNA synthetase.
GN   YARS.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney marrow;
RA   Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,
RA   Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zon L.I.,
RA   Kanki J.P., Look A.T., Chen Z.;
RT   "Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue.";
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY398430; AAQ97863.1; -.
KW   Aminoacyl-tRNA synthetase.
SQ   SEQUENCE   529 AA;  59505 MW;  C3314003813B9336 CRC64;

  Query Match             28.8%;  Score 453.5;  DB 2;  Length 529;
  Best Local Similarity   39.8%;  Pred. No. 5.4e-19;
  Matches  123;  Conservative   43;  Mismatches   78;  Indels   65;  Gaps   14;

Qy         44 EEKKLRV--ENAKLKKEIEELKQELIQAEIQNG----VKQIAFPSGTPLHANSMVS---- 93
              || |: :  :|  :||::::   |    |  ||    || : |    |||:  ::
Db        228 EESKIDLLDKNQDVKKKLKKAFCEPGNVE-NNGVLSFVKHVLF----PLHSEFVIKRDPK 282

Qy         94 ----------ENVIQSTAVTTVSSGTKEQIKGGTGDEK--------KAKEKIEKKGEKKE 135
                       | | :  |         |||    || |          |  : | || |  |
Db        283 FGGDKVYTDFEEVEKDFAA--------EQIH--PGDLKASVELALNKLLDPIRKKFESPE 332

Qy        136 -KKQQSIA--------GSADSKP--------IDVSRLDLRIGCIITARKHPDADSLYVEE 178
               ||  | |         | |   |         :  ||||:|:| :|:  |||||||||:|:
Db        333 LKKLTSSAYPEPSKNKGGAKGNPKQTTDDDEVIPSRLDIRVGKVISVEKHPDADSLYLEK 392

Qy        179 VDVGEIAPRTVVSGLVNHVPLEQMQNRMVILLCNLKPAKMRGVLSQAMVMCAS---SPEK 235
              :|||| |||||||| ::  ||:|:|:|:|||||| ||||: |||||:|||    | |
Db        393 IDVGEEQPRTVVSGLVAYITEEQLQDRLVVLLCNLKPQKMRGIESQAMVLCASIEGEPRK 452

Qy        236 IEILAPPNGSVPGDRITFDAF-PGEPDKELNPKKKIWEQIQPDLHTNDECVATYKGVPFE 294
              :| | || ||  |||:  : :  |:|| || ||||::|::| ||  : | || :|
Db        453 VEPLDPPEGSAAGDRVYVEGYESGKPDDELKPKKKVFEKLQVDLKISGEFVAQWKEQNLM 512

Qy        295 VK-GKGVCR 302
               | |:  |:
Db        513 TKLGRITCK 521


Search completed: January 14, 2005, 15:58:24
Job time : 193 secs
```